INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

GENERATION AND REFINEMENT OF A CONTINENTAL SCALE DIGITAL ELEVATION MODEL BY INTEGRATING CARTOGRAPHIC AND REMOTELY SENSED DATA: A GIS-BASED APPROACH

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Philosophy in the Graduate

School of The Ohio State University

By

Hongxing Liu, B.S. M.S.

*****

The Ohio State University

1999

Dissertation Committee:

Professor Morton E. O'Kelly, Adviser

Professor Kenneth C. Jezek

Professor Carolyn J. Merry

Approved by:

Adviser

Graduate Program in Geography

UMI Number: 9931639

UMI Microform 9931639
Copyright 1999, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized
copying under Title 17, United States Code.

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

ABSTRACT

Digital Elevation Models (DEMs) are of fundamental importance to a wide range of geo-scientific and environmental studies. This research addresses technical and theoretical issues involved in the construction of a large-scale DEM with multiple data sources. Enormous data volume, diverse data sources and different terrain types complicate the large-scale DEM generation process. To tackle technical difficulties and challenges, several new methods have been developed and some existing algorithms have been improved. Having these newly developed and extended methods and algorithms along with available GIS functions, we constructed a complete, consistent, and high-resolution DEM over the Antarctic by integrating a variety of cartographic and remotely sensed data.

This research demonstrates that the GIS technology represents an effective means of integrating not only various types of topographical data, but also a pool of alternative computational methods and techniques for the creation of large-scale DEMs. Nevertheless, currently available GIS software did not provide an entire solution to the development and refinement of large-scale DEMs.

In this research, a number of error detection techniques have been developed for efficiently checking and eliminating gross errors in source data. New interpolation procedures were designed for gridding traverse data and contour-based topographical

data. The determination of an optimal interpolation spacing was pursued. A conceptual framework was proposed for integrating and fusing complementary and competitive topographical data. In the case of the Antarctic, a set of criteria were developed for the selection of competitive data to exploit the best elements of available sources. It is also demonstrated that spatially and functionally complementary topographical data sources can be synergistically combined during the interpolation stage to create a more reliable and accurate elevation model than what a single source could create alone. A new algorithm was developed for merging individual DEM data sets into a seamless DEM mosaic along irregular lines. For DEM refinement, the potential of the state-of-the-art remote sensing techniques, including shape-from-shading, SAR stereo and interferometric SAR (InSAR), are discussed and demonstrated. A technical strategy is formulated for future improvement of the Antarctic DEM.

The accuracy and error pattern of the resultant Antarctic DEM were evaluated in a number of ways. Comparisons with high-resolution satellite images at different scales demonstrate that the DEM not only describes the broad topographical configuration and true ice drainage pattern, but also captures unprecedented detail of the terrain surface. We believe that this DEM product represents the most detailed and accurate digital description of the Antarctic topography published to date.

Dedicated to my mother and father

ACKNOWLEGMENTS

At the time of completing this dissertation, I feel indebted to many people for both offering direct advice and comments on the manuscript and for providing various assistance to support this research.

First of all, I wish to express my sincere gratitude to Professors Morton E. O'Kelly and Kenneth C. Jezek for their unfailing advice and valuable instructions at all points along the way of my doctoral study. Their expertise, insights and passion for research, and their enlightening lectures have strongly influenced my research direction and stimulated my interests in GIS, satellite remote sensing and spatial data analysis in general and digital elevation models in particular. During the preparation of the dissertation, their critical and thoughtful comments and suggestions gave me great help in clarifying the concepts, sharpening the thinking and broadening the vision, and also made the dissertation research more challenging and fruitful.

Special thanks are extended to Dr. Carolyn Merry. While serving on my dissertation committee, she carefully reviewed my research proposal and the draft manuscript and offered constructive comments. In addition, I benefit enormously from taking her informative classes in remote sensing. My thanks also go to Dr. Ellen Mosley-Thompson for her comments on my research proposal and her lecture in global environmental issues.

Also, I would like to acknowledge other professors who taught me. In particular, I wish to express my appreciation to a series of cartography classes offered by Dr. Harold Moellering, quantitative methods classes by Dr. Randy Jackson and Dr. Steven Holloway, and digital photogrammetry and GPS classes offered by the Department of Geodetic Science.

In addition, I want to note my gratitude to Marty Marra for his instructions on the use of the Vexcel SAR Stereo Toolkit and Rick Guritz for his advice on the use of ASF InSAR software.

I owe special thanks to my colleagues in the Remote Sensing Lab at the Byrd Polar Research Center for their friendship and constant help. I am particularly grateful to Katy Noltimier for spending a lot of time to review and edit my academic writings, Hong-Gyoo Sohn for generously sharing his knowledge in photogrammetry, various satellite data and computer programming, Zhiyuan Zhao for helping to produce the SAR simulation image, and to Beata Csatho, Daniel Roman, Rick Forster and Francois Baumgartner for useful discussions.

Finally, I want to record my deep gratitude to my wife and colleague, Biyan Li, for her emotional encouragement and great aid in data processing and illustration preparation, to my parents for whom I could not fail and for whom this dissertation is dedicated to.

This research was supported by grants from NASA's Pathfinder and Polar Oceans and Ice Sheets Programs.

VITA

October, 1961.....Born – Xingjiang, China

1982.....B.S. Geography, Shaanxi Normal University

1985.....M.S. Geograpgy, Peking (Beijing) University

1994.....M.S Urban Studies, The University of Sydney

1986-1992.....Lecturer, Dept. of Geography
Peking (Beijing) University, China

1994-1995.....Technician Assistant, Dept. of Geography
The Ohio State University

1996-present.....Graduate Research Associate
Remote Sensing Lab
Byrd Polar Research Center
The Ohio State University

PUBLICATIONS

Refereed Research Publications

1. Liu, H. and K.C. Jezek (1999). "Investgating DEM error patterns by directional variograms and Fourier analysis." *Geographical Analysis*, Vol. 31, No. 3.

2. Liu, H., K. Jezek and B. Li, "Development of Antarctic digital elevation model by integrating cartographic and remotely sensed data: a GIS-based approach," submitted to *Journal of Geophysical Research* (accepted).

3. Liu, H., and B. Li (1992). "Application of shift-share methods on analysis of urban economic structure," *Territorial Development and Management*, Vol.2, No.2, p44-51 (in Chinese).

4. Dong, L. and H. Liu (1991). "Development and planning concepts for the great urban agglomeration in the central Liaoning province," *City Planning Review*, No. 5, p21-24 (in Chinese).

5. Liu, H. (1988). "On location of enterprises and development of town," *City Planning Review*, No.4, p6-8 (in Chinese).

6. Liu, H. (1988). "Systems analysis and regionalization of population capacity in the Southern Ningxia," Economic Geography, Vol. 8., No.4, p249-252 (in Chinese).

7. Liu, H. (1987). "Analysis of urbanization characteristics and trend in Wenzhou," City Planning Review, No. 2, p39-44 (in Chinese).

FIELDS OF STUDY

Major Field: Geography

TABLE OF CONTENTS

Page

Abstract.....ii

Dedication.....iv

Acknowledgments.....v

Vita.....vii

Publications.....vii

Field of study.....viii

Table of contents.....ix

List of Tables.....xii

List of Figures.....xiii

Chapters

1. Introduction.....1

   1.1 Background.....1

   1.2 Research scope and objectives.....3

   1.3 Organization of the research.....5

2. Large-scale DEM generation and GIS technology.....8

   2.1 Digital representation of a terrain surface.....8

   2.2 Large-scale DEM generation process.....10

2.3 GIS technology in large-scale DEM generation.....15

3. Consistent and Error-free Topographical Database.....22

3.1 Topographical database.....22

3.2 Characteristics of different data acquisition techniques.....22

3.3 Topographical data sources for Antarctica.....27

3.4 Planimetric and vertical reference systems.....34

3.5 Error detection and correction techniques.....39

3.5.1 Errors in topographical data.....39

3.5.2 Global statistical analysis and cross-validation methods.....40

3.5.3 Scientific visualization and image simulation methods.....42

3.5.4 Locally adaptive and robust statistical analysis method.....51

3.6 Summary.....62

4. Spatial interpolation of topographical data.....65

4.1 Overview of interpolation algorithms.....65

4.2 Interpolation of evenly distributed point data.....69

4.3 Interpolation of traverse data.....78

4.4 Interpolation of contour-based data.....88

4.5 Determination of optimal interpolation interval.....100

4.6 Summary.....104

5. Data integration and merging.....107

5.1 Conceptual framework for data integration.....107

5.2 Integration of complementary source data.....109

5.3 Selection of competitive source data.....112

5.4 Seamlessly merging of individual DEM data sets....117
6. Continental Scale DEM and Quality Assessment....123
6.1 Seamless continental scale DEM....123
6.2 Quality assessment....133
6.2.1 Comprehensive view of quality....133
6.2.2 Horizontal resolution and positional accuracy....134
6.2.3 Reliability and vertical accuracy....135
6.3 Error pattern analysis....140
7. DEM Refinement by Spaceborne Techniques....155
7.1 Improvement of large-scale DEM....155
7.2 Shape-from-shading technique....157
7.2.1 Overview of shape-from-shading technique....157
7.2.2 Algorithm implementations with application to SPOT image....159
7.3 SAR stereo technique....171
7.3.1 Principles of SAR stereo technique....171
7.3.2 SAR stereo processing with Radarsat data....176
7.4 Interferometric SAR technique....185
7.4.1 Theoretic background of SAR interferometry....185
7.4.2 Elevation data extraction with ERS Tandem data....191
7.5 Technical Strategy and Terrain Characteristics....203
8. Conclusions....209
List of References....216

LIST OF TABLES

Table Page

6.1 Accuracy assessment of the DEM.....139

7.1 Suitability of different spaceborne techniques in terms of terrain types.....206

LIST OF FIGURES

Figure		Page
2.1	Diagram of a large scale DEM generation procedure	14
2.2	Graphical User's Interface (GUI) for the IGIAS software	17
2.3	Graphical User's Interface (GUI) for the SpaceView software	20
3.1	Scale and coverage of cartographic data	28
3.2	Distribution of ground survey and GPS data	33
3.3	Relation of topography, geoid and ellipsoid	37
3.4	Geoid undulation of the Antarctic	38
3.5	Geometric definition of local incidence angle	45
3.6	A mislabeled contour detected in hill-shaded image through synthetic stereo view	47
3.7	Detection of DEM positional errors using SAR image simulation technique	50
3.8	Definition of octant neighbor points	53
3.9	Super-block based sorting and spiral searching	54
3.10	A typical neighborhood configuration	56
3.11	Erroneous data points in the Evans RES data set detected using a locally adaptive and robust statistical analysis method	63
4.1	Spatial distribution pattern of preprocessed ERS-1 Radar altimeter data	

in the Victoria Land, Antarctica.....70

4.2 Interpolation result with Universal Kriging on the evenly distributed data set.....74

4.3 Interpolation result using minimum curvature spline with regularized option on the evenly distributed data set.....77

4.4 Interpolation result with TIN linear option on the evenly distributed data set.....79

4.5 Interpolation result with TIN quintic option on the evenly distributed data set...80

4.6 Spatial distribution pattern of airborne traverse data over the Evans Ice Stream, Antarctica.....82

4.7 Interpolation result with universal Kriging on the traverse data set.....83

4.8 Interpolation result with minimum curvature spline on the traverse data set.....84

4.9 Interpolation result with TIN linear option on the traverse data set.....86

4.10 Interpolation result with the combination of quadrant neighborhood based IDW and TIN quintic on the traverse data set.....89

4.11 Interpolation result with TIN linear option on the contour data.....91

4.12 Interpolation result with TIN quintic option on the contour data.....92

4.13 Interpolation result with linear gradient descent on the contour data.....94

4.14 Interpolation result with TOPOGRID on the contour data.....97

4.15 Interpolation result with modified two-stage TOPOGRID procedure on the contour data.....99

4.16 Radial power spectrum of mountainous terrain.....103

4.17 Changes in average height error with interpolation grid spacing.....105

5.1 Integration of various input data sources for DEM creation in Antarctic Peninsula.....111

5.2 Distribution of input data sources used in the final Antarctic DEM.....114

5.3 Data merging scheme using cubic Hermite blending function.....119

5.4 Comparison of the cubic Hermite blending function-based merging method with conventional method.....121

6.1 Shaded relief image of the DEM at continental scale.....124

6.2 Ice drainage pattern and ice flow lines derived from Antarctic DEM.....126

6.3 Comparison of the DEM shaded relief image with Landsat TM image in OSCAR II Coast of Antarctic Peninsula.....128

6.4 Comparison of the DEM shaded relief image with DISP satellite photograph in Transantarctic Mountains.....129

6.5 Comparison of the DEM shaded relief image with Radarsat SAR Quicklook image mosaic in Victorial Land of Antarctica.....130

6.6 Perspective view of a shaded relief image draped on DEM around Anverse Island and Brabant Island of Antarctic Peninsula.....131

6.7 Perspective view of Radarsat SAR image draped on the DEM in Ponnell Coast, Antarctica.....132

6.8 Accuracy assessment of the DEM by comparing with GPS, RES and leveling data.....138

6.9 Comparison of McMurdo Dry Valley DEM data with reference data.....142

6.10 Spatial distribution of absolute errors.....143

6.11 Directional variograms and covariograms of DEM errors.....146

6.12 Spatial autocorrelation structure of the DEM errors.....150

6.13 Power spectral density profiles of the terrain surface.....152

7.1 SPOT image and shape-from-shading derived DEM over the Crary Ice Rise, Antarctica.....160

7.2 Elevation data derived from local shape-from-shading technique.....164

7.3 Elevation data derived from global shape-from-shading technique.....169

7.4 Contours and perspective view of DEM derived from global shape-from-shading algorithm.....170

7.5 Sensor state vector determination and range/Doppler intersection in SAR stereo technique.....174

7.6 Stereo pair of RADARSAT SAR data over Terra Nova Bay area.....178

7.7 Distribution of matched points from SAR stereo processing.....180

7.8 Noisy surface directly derived from Vexcel stereo tool.....182

7.9 Comparison of hill-shaded DEMs derived from SAR stereo technique and topographic map.....183

7.10 Comparison of contours derived from SAR stereo with USGS 1:250,000 topographic map.....184

7.11 Geometry of cross-track interferometric SAR.....187

7.12 Amplitude and correlation maps of InSAR pair over Dry Vally, Antarctica.....192

7.13 Flattened (Deramped) phase of interferogram.....196

7.14 Hill-shaded DEM derived from InSAR technique.....200

7.15 Hill-shaded DEM derived from the combination of 1:50,000 and 1:250,000 USGS topographic maps.....201

7.16 Accuracy assessment of InSAR derived elevation.....202

CHAPTER 1

INTRODUCTION

1.1 Background

Generally speaking, a Digital Elevation Model (DEM) can be any digital cartographic representation of the terrain surface (Burrough, 1986). In this research, the term 'Digital Elevation Model' (DEM) is used in a narrow sense to refer to a two-dimensional array of altitude values at regularly spaced ground positions. The term 'Digital Terrain Model' is also interchangeably used in the literature to mean the same thing (Burrough, 1986; Weibel and Heller, 1991).

Digital elevation models (DEMs) are used in a wide range of fields (Burrough, 1986; Gao, 1995). Various terrain parameters can be derived from DEMs, such as slope, aspect and surface curvature. These parameters provide an important basis for many geo-scientific and environmental applications. Examples include estimating soil erosion, water run-off and potential flooding areas, analyzing glacier dynamics, and delineating drainage patterns, to name just a few. In civil engineering, high-resolution DEMs are often utilized to solve cut-and-fill problems and select locations for dams and roads. In the fields of cartography and remote sensing, DEMs are widely used in the automated

topographical mapping, image ortho-rectification and scientific visualization fields (Jensen, 1995).

The past several decades have witnessed an expansion in DEM related research. In addition to conventional tacheometric ground survey and aerial photogrammetry methods, many new instruments and techniques have emerged for capturing topographical data, including global positioning system (GPS), satellite image based stereo photogrammetry, interferometric SAR (InSAR), radar altimetry and laser altimetry. Meanwhile, substantial progress has been achieved in the theories, methods and algorithms regarding the construction of DEMs. However, a brief review of the literature reveals that most previous DEM studies deal with situations concerned with only a single data acquisition technique or a single data source for a relatively small area. The issues involved in the development of a large scale DEM with multiple data sources and techniques are rarely discussed in the literature. Consequently, the knowledge and information is still sparse and inadequate in this regard.

Coverage requirements for DEMs have been extended from local to regional, from regional to national, continental and even global scale due largely to their importance in national and international mapping projects and global environmental change studies. Though similar in many aspects to local DEM generation, the construction of a large-scale, high resolution DEM is complicated by three major facts. First, the volume of both input and output data is enormous. Second, diverse data sources are often utilized. The data type, format, reference systems, the density, spatial distribution pattern, the accuracy and the original measuring instrument and sensor may vary from source to source. Third, various terrain types might be included due to the large areal coverage. The combination

of all these factors raises a spectrum of new issues and questions. For instance, what principles and techniques should we use to integrate and fuse a variety of data sources so as to produce a coherent high-quality DEM over a large area? How do we design and adapt the techniques and algorithms for various types of source data and different types of terrain? In the face of a large volume of data and an extensive ground area, how can we reduce the computational cost and automate the DEM generation process? What data acquisition techniques are available and appropriate for updating and enhancing a large scale DEM? What are their comparative advantages and limitations? These questions form the basis for this study.

1.2 Research scope and objectives

This study addresses the technical and theoretical issues encountered in large-scale DEM generation and explores the potential of the Geographical Information Systems (GIS) technology in tackling these issues. The Antarctic continent is taken as the case study area, based on the following reasons:

- A high-resolution DEM over the entire Antarctic is required for correcting terrain distortions of SAR imagery in the Radarsat Antarctic Mapping Project (RAMP), which has an ultimate goal of producing the first-ever, complete, and seamless orthorectified SAR image mosaic of the Antarctic.
- As will be described later, a variety of topographical data and satellite image data are available and accessible for the author to address the research questions.
- Various types of topography exist for the Antarctic continent. This ensures that the research findings are representative of and applicable to other regions or continents.

- The hostile climate, inaccessible environment and extremely high latitude of the Antarctic continent impose serious difficulties and challenges for conventional data acquisition methods and call for state-of-the-art remote sensing techniques. On the other hand, virtually no vegetation coverage in Antarctica allows for extracting the surface elevation data from the remote sensing data without the need of correcting for the vegetation canopy effect.

With the Antarctic continent as the study case, the specific objectives of this study include:

1) To investigate and demonstrate the capability of GIS in integrating different information sources and different techniques and algorithms, and in automating the large-scale DEM generation process;
2) To identify the techniques for converting different types of source data into a consistent topographical database and develop the techniques for detecting and correcting errors in the source data to make the topographical database free of gross errors and blunders;
3) To design and evaluate spatial interpolation algorithms for different types of data sources under different terrain conditions and develop the techniques for the determination of an optimal interpolation interval;
4) To formulate a conceptual framework for the integration and fusion of various topographical data sets, to establish a set of criteria for selecting competitive topographical data sources, and to develop the technique for seamlessly merging individual DEM data sets.

5) To explore and demonstrate the capability of advanced spaceborne techniques in updating and refining a large-scale DEM, and to examine their suitability and feasibility for different types of terrain; and

6) To make a quality assessment of the final continental scale DEM in terms of horizontal resolution, vertical accuracy and reliability.

Overall, this research develops a comprehensive and coherent strategy for integrating a variety of topographical data sources and related techniques and algorithms to construct a large-scale DEM of the best possible quality in a GIS environment.

1.3 Organization of the research

The research background and objectives have been presented in the preceding sections. Chapter 2 starts with an overview of different digital surface representation methods. The strengths and weaknesses of raster-based DEM are discussed in comparison with other data models. Then the component steps required for the construction of a continental scale DEM are described. This is followed by a discussion of the capabilities of GIS in supporting a large-scale DEM generation process. The GIS software packages constructed to assist this research are briefly introduced.

Chapter 3 examines the technical issues of how to compile a consistent and error-free topographical database, which is the first step toward data integration and fusion. This chapter proceeds from a brief description of available topographical data sources and corresponding data acquisition techniques for the case of the Antarctic continent. The comparative advantages and limitations of each source are discussed in the light of measuring accuracy, the sampling density and pattern, ground coverage, acquisition cost

and data availability. Then the need for a common planimetric coordinate system and vertical datum is highlighted in the context of data adjustment and conversion. The final section presents a number of error detection and correction techniques designed for different types of source data, with an emphasis on a newly developed error detection algorithm for irregularly distributed data sets.

In Chapter 4 the performances of different spatial interpolation methods are examined in terms of the type and distributional pattern of input data. The focus is placed on two interpolation procedures specially designed for traverse data and contour-based data. The determination of an optimal interpolation interval is addressed in the final section of this chapter.

Chapter 5 places the focus on the integration and fusion of various topographical data sources. At the start of the chapter, a conceptual framework is established and different integration scenarios are presented in terms of the functional and spatial relationships between different data sets. Then, for each scenario the integration strategy and technique are discussed with examples. A new technique is put forward for seamlessly merging individual DEM data sets.

Chapter 6 presents the resultant Antarctic DEM and assesses its horizontal resolution and vertical accuracy. In comparison with real satellite images at different scales, the unprecedented topographical details present in our DEM are demonstrated. To check the topological and geomorphological consistency of the DEM, the ice drainage pattern is derived over the Antarctic continent and compared with the satellite image mosaic and the published drainage map. To assess the absolute vertical accuracy, the DEM is

compared with GPS, airborne radar data and geodetic leveling traverses. The error pattern is also analyzed using directional variograms and Fourier analysis.

In Chapter 7, a number of advanced spaceborne remote sensing techniques are examined for the purpose of enhancing DEM resolution and accuracy. Two shape-from-shading algorithms are implemented and tested on a SPOT panchromatic image at Crary Ice Rise. The SAR stereo technique is applied to a Radarsat SAR stereo pair in the Terra Nova Bay area. Interferometric SAR analysis is performed on ERS Tandem SAR data in the Dry Valley region. The performances and applicability of these techniques are discussed in terms of morphological and textural properties of terrain surface, and a technical strategy of using spaceborne techniques to enhance a large-scale DEM is formulated in the case of the Antarctic.

The final chapter summarizes the research findings and presents conclusions and implications derived from this study.

CHAPTER 2

LARGE-SCALE DEM GENERATION AND GIS TECHNOLOGY

2.1 Digital representation of a terrain surface

Traditionally, a terrain surface is depicted by analog methods, such as contour-based topographic maps, relief models, photographs and sketches. With the rapid advance in computer technology, various digital methods have been developed to collect, store and convey terrain surface information. In terms of fundamental data structure, digital terrain representation methods can be classified into three categories: raster-based DEM methods, vector-based methods, and mathematical models. A brief review of each method is provided below to highlight the strengths and weaknesses of the raster-based DEM in a comparative context.

A mathematical model assumes that the terrain surface is continuous and single-valued. Under this assumption, a complete surface is often split into square cells or irregularly spaced patches of roughly equal area and the surface functions are fitted based on the point observations within each patch. Fourier series functions are often used for modeling the global trend of the surface, while multi-polynomial functions are used for fitting local patches. The mathematical model is the most compact and elegant format, but a large number of coefficients of polynomial or Fourier series functions have to be

calibrated and stored, and the elevation values at given locations have to be evaluated from the functions when they are needed. In practice, this method is seldom used, except for modeling the general and global trend of the terrain (Burrough, 1986).

Vector-based methods include digitized contour lines, spot points, elevation profiles, and a Triangulated Irregular Network (TIN). Because contours are widely available on existing topographical maps, they can be easily converted to digital format by automated scanning or manual digitizing. The digitized contours and spot points, stored in vector format, is the graphical equivalent to a topographic map. Elevation profiles are usually derived from the survey trajectories of expeditions or flight traverses. A TIN is a terrain model consisting of a set of connected triangular facets, usually based on a Delaunay triangulation of irregularly spaced elevation points of observation. These vector-based methods have the advantages of having more precise geometric locations of measurements and less data redundancy. They are also able to adapt the data density according to terrain conditions and explicitly represent the topological structure of the terrain surface. For instance, a TIN model can vary the spacing of network node points for gathering the extra information in areas of complex terrain, while avoiding the huge amount of redundant data in areas of simple and flat relief. Topological features like terrain break lines (ridge, valleys, and streams) and critical points (peaks and pits) can be easily incorporated into a TIN model. Despite the advantages described above, the practical application of vector-based methods is seriously limited by the inefficiency for data retrieval, and the difficulties in incorporating vector data into quantitative analysis and mathematical modeling.

A raster-based DEM uses a regularly spaced grid of elevations to represent the terrain surface. This method has its inherent weaknesses, including the large volume of data, especially data redundancy in flat areas, difficulty in adapting to the terrain complexity without changing the grid cell size, and the inability of explicitly representing the topological structures. The advantages of the raster-based DEM lie in the simple data structure for programming and machine storage, the strong capability of representing the continuous surface, the efficiency in quantitative analysis and cartographic modeling, and the compatibility to the remotely sensed data. Given these advantages, the raster-based DEM has become the most commonly used digital format for representing terrain surface.

2.2 Large-scale DEM generation process

A terrain surface consists of an infinite number of points, but only a limited number of points can be measured in practice. With a representative and meaningful sample of elevation measurements, the goal of DEM generation is to computationally derive a structured and regularly spaced elevation grid for representing the essence of the topography.

To achieve complete spatial coverage and better quality data, multiple topographical data sources are often collected and assembled in the creation of a DEM over an extensive area, especially at a continental or global scale. The collected source data may originate from different acquisition instruments and techniques. As a result, the accuracy, density, spatial pattern, and coverage often vary from source to source. Topographical data commonly used in practice include direct measurements from survey

and remote sensing techniques and cartographic data digitized from topographical maps. In addition, terrain morphological features are often incorporated as surface structural constraints in the DEM computation. Previous research suggests that the terrain behavior and shape are mainly controlled by the critical topological features like ridges, valleys, streams, coastlines, lake shorelines, peaks, pits, passes and pales (Brassel and Weibel, 1988; Weibel, 1992). This is known as the Warntz network in analytical cartography (Warntz, 1966; Mark, 1979; Wilcox and Moellering, 1995). These structural elements have more topographical information content than any other ordinary points on the terrain surface. Therefore, the inclusion of topographical structures as input data is an efficient way to improve the DEM quality.

Most often, different topographical data sets are measured and processed using different geographic reference systems. Before loading data into the DEM generation model, input data sources must be geocoded and co-registered using a common planimetric coordinate system and vertical datum. In addition, the source data might be contaminated by various errors, and thus error detection and correction operations are normally required. Compiling and constructing a consistent and error-free topographical database represents a critical step for large-scale DEM generation.

In most cases original topographical measurements from ground survey and remote sensing techniques are irregularly distributed over space in either a scattered or traverse pattern. They are often conveniently recorded in the vector format as a series of xyz triplets. Cartographic data digitized from topographical maps are dominated by vector lines (contours). Morphological features can be also described by points (local maxima or minima) and lines (ridges and valleys).

The principal computation involved in DEM generation is to interpolate the vector dominated topographical source data into a regular elevation grid based on smoothness and continuity assumptions of the terrain surface. Although many spatial interpolation algorithms are available, their performance varies wildly according to the format, density and pattern of the input data. Selecting and designing an appropriate interpolation algorithm for each type of source data is another requirement for large-scale DEM generation, in contrast to a local DEM generation where only one interpolation model is needed to approximate a small terrain surface.

When several source data sets overlap over space, we need to develop rules and criteria either to select the best data set or to take the best components of each data set and combine them together during the interpolation. In some cases, a number of coincident and redundant data sets emerge from the repeated observations of the same area using a specific instrument or from slightly different processing techniques on the same set of raw data. Reducing data redundancy while increasing data reliability is the goal for integrating this type of data.

Due to the tremendous amount of data and the expanse of ground area involved in large-scale DEM generation, it is computationally impossible to create a complete DEM grid at one run. The common practice is to create DEM grids block by block. Correspondingly, the source data have to be clipped or merged according to the tiling of DEM blocks. Discrepancies between individual DEM blocks often occur due to edge effects of interpolation or inconsistencies of the source data used in different blocks. Minimizing the discontinuity and seamlessly merging the adjacent DEM blocks is another aspect of data integration. In addition, when individual DEM grids have different

grid spacing, resampling algorithms are required to densify or coarsen some of the DEM blocks before the merging operation is applied.

At the final stage, the resolution, accuracy and reliability of the resulting DEM mosaic need to be evaluated. Since the DEM mosaic is created based on the existing topographical data, the whole DEM grid, especially the areas with poor quality, needs to be updated and enhanced when new topographical data sources become available.

In summary, a large scale DEM generation process technically consists of the following component steps (Figure 2.1):

- Collecting and assembling available topographical data sources, and converting them into coherent digital formats, vector or raster where appropriate;
- Projecting and adjusting various data sources using a common planimetric and vertical reference system;
- Checking and editing the source data to remove errors;
- Clipping and merging source data into a number of overlapping blocks;
- Designing interpolation algorithms for different types of source data, and interpolating source data into DEM blocks using an appropriate interpolation scheme and an optimally determined grid spacing.
- Based on prior knowledge of the relationship between and comparative advantages of source data, performing data selection and fusion during interpolation;
- Adjusting the grid spacing of DEM blocks using grid-to-grid resampling methods; Merging the adjacent DEM blocks into a complete and seamless DEM mosaic;
- Assessing the quality of the resultant DEM and conducting error analysis; and

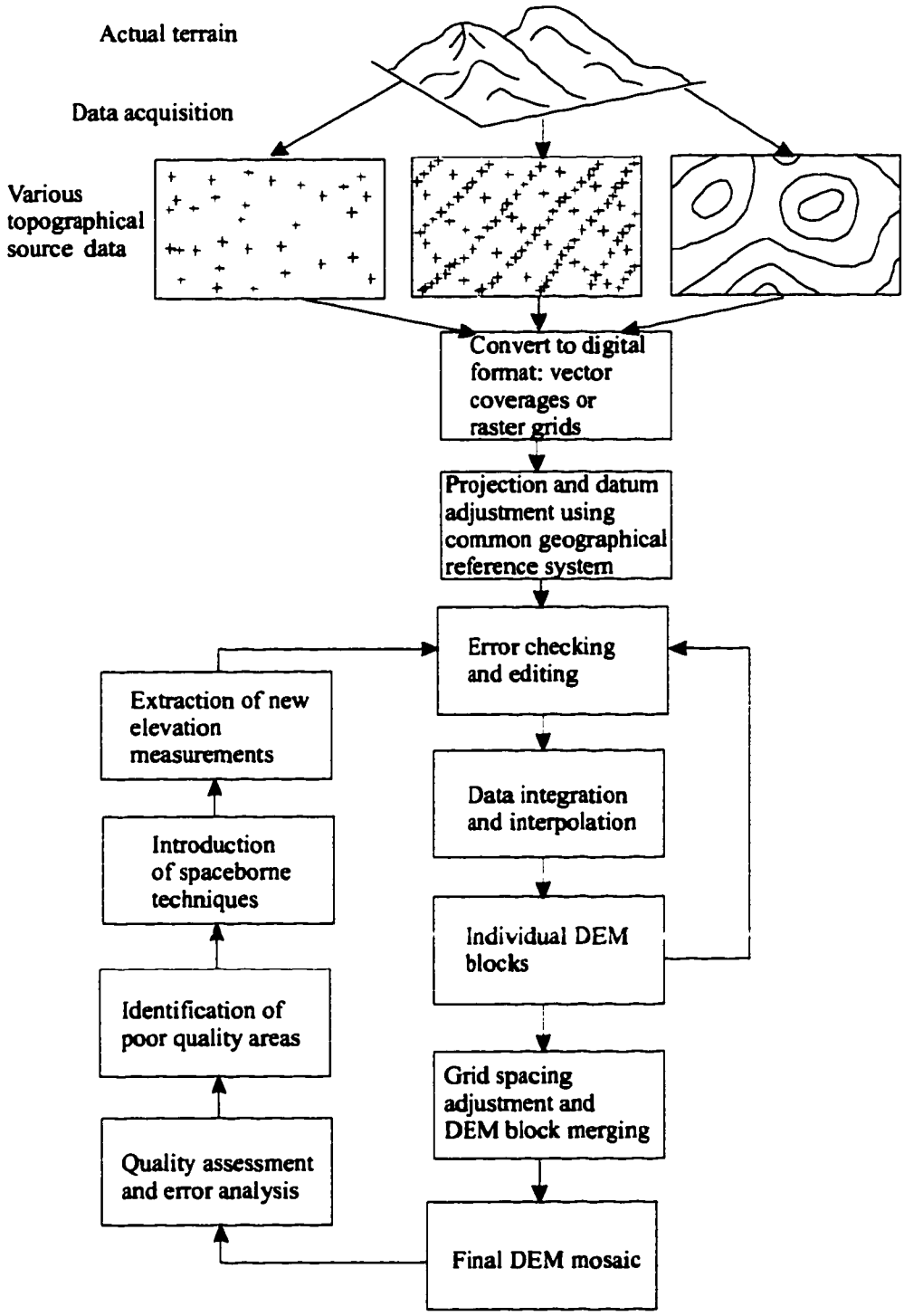


Fig. 2.1: Diagram of a large scale DEM generation procedure

- Identifying the poor data quality areas and formulating a plan for updating and enhancing the DEM.

2.3 GIS technology in large-scale DEM generation

From the preceding discussion, it is quite clear that the generation of large-scale DEMs needs not only additional pre-processing and post-processing steps, but also a pool of flexible and alternative computational algorithms for integrating and modeling diverse topographical data. The conventional stand-alone software system used to support data acquisition and computational derivation of a small scale DEM with a single data source is no longer sufficient for this complex situation. A comprehensive software system is required that is capable of combining multiple topographical data types, providing for their storage, retrieval, and display, and incorporating a wide range of relevant data manipulation and modeling functions. In this research, the Geographical Information System (GIS) technology is used to construct such an integrated software system.

The GIS technology developed from the need to handle and manipulate geographically referenced spatial data. The ability of GIS to integrate diverse information is well known and frequently cited as its major source of power and appeal. In the inception stage, the major functionality of GIS lies in the storage, retrieval, manipulation and display of geographic data. Since the late 1980s, analytical and modeling capabilities have been substantially improved by introducing the concept of map algebra (Tomlin, 1990), spatial statistical analysis and mathematical modeling methods (Goodchild et al., 1992; Bailey, 1994; O'Kelly, 1994; Haining, 1994). These emerging capabilities lay a firm foundation

for the GIS technology to satisfy the functional requirements of a large-scale DEM generation.

Based on internal analytical and graphical functions of the proprietary GIS software ARC/INFO and external functions added by C language programming, a software package, Integrated GIS and Image Analysis System (IGIAS), was developed on Unix workstations to support the large-scale DEM generation process in this study. The graphical user interface (GUI) of the IGIAS was built by using AML (Arc Macro Language), X-window library and Motif (Figure 2.2). Under this common GUI, the component functions are accommodated and the principal DEM generation steps are linked together in a transparent and seamless manner. Users have an easy access to both topographical databases and the analytical and modeling functions required for large-scale DEM generation.

Within the IGIAS, various import functions are available for converting topographical data sources of diverse formats into coherent and topologically structured ARC/INFO vector coverages (point, line, polygon) or generic raster layer (grids), generic in the sense that the grid data structure can accommodate any type of numerical values, positive or negative, integer or float, single or double precision. Remotely sensed image data can also be easily incorporated for DEM quality verification and surface structural feature extraction. In this architecture, both vector and raster topographical data can be efficiently processed and integrated to model the terrain surface.

Map projection, datum conversion between geoid and ellipsoidal models, and a number of interactive or automatic error detection and editing routines are developed and embedded in the IGIAS. With these routines, various topographical data coverages and

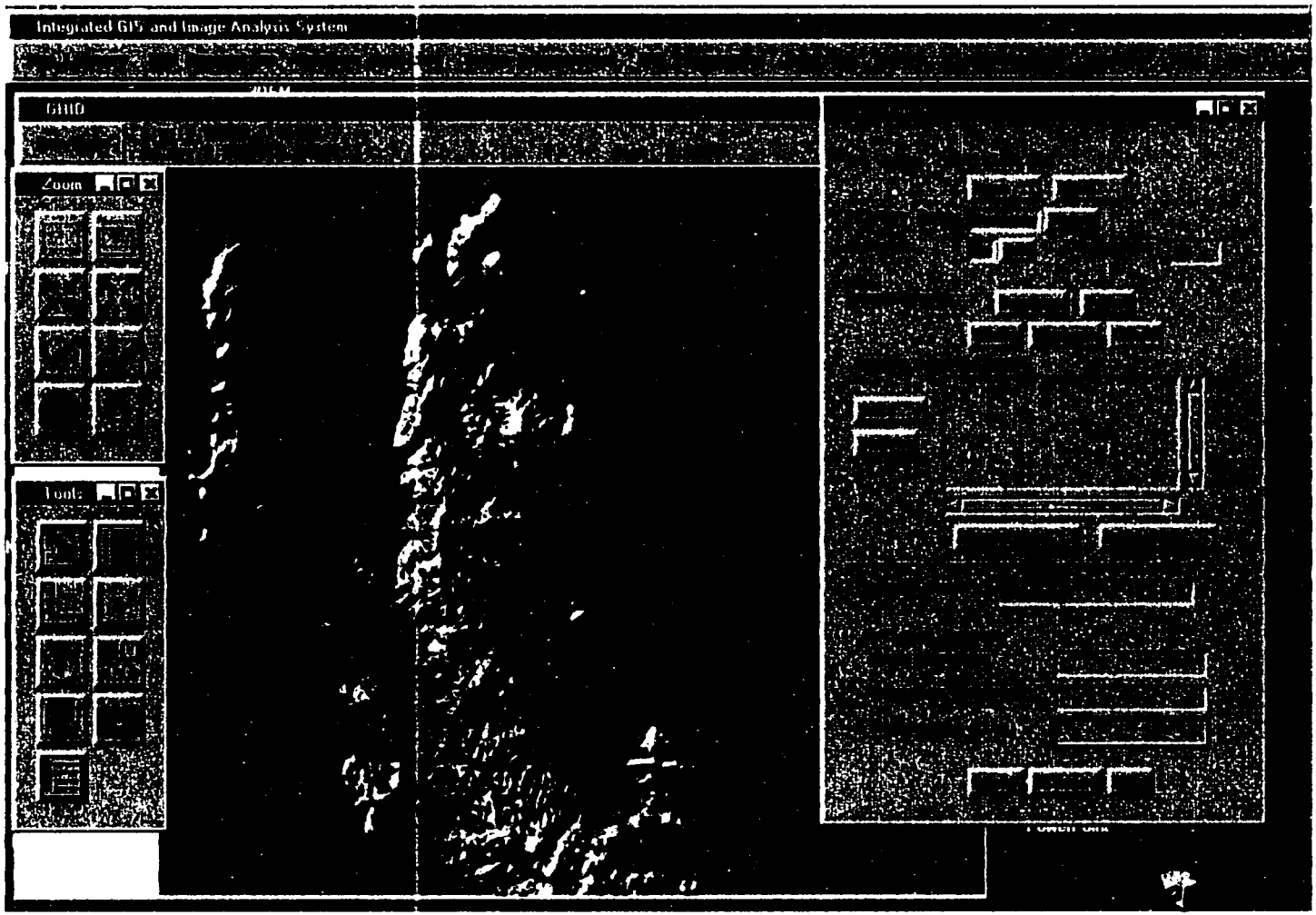


Fig. 2.2: Graphical User's Interface (GUI) for the IGIAS software

layers can be spatially correlated with each other to form a consistent and reliable topographical database in support of the DEM computation. The data elements of topographical coverages and layers can be selected in terms of either logical expression of associated attributes or by graphically defined regions of various shapes. The properties of data elements or the whole data layers can be interactively queried, and summary statistics can be obtained. Functions for clipping and merging different topographical data coverages are also included in the IGIAS. These spatial data manipulation capabilities have greatly facilitated topographical database compilation, management and updating.

Various spatial interpolation algorithms have been incorporated and developed in the IGIAS, including a TOPOGRID algorithm, TIN-based linear and quintic algorithms, quadrant or octant neighborhood based IDW (Inverse Distance Weighted) algorithm, minimum curvature spline, ordinary and universal Kriging, and steepest descent linear interpolation algorithm. The combination of logical overlay and grid-based map algebra functions furnishes the data fusion requirement after interpolation. A cubic Hermite blending weight algorithm is developed to seamlessly merge DEM grid blocks along an irregular buffer zone. 2D FFT (Fast Fourier Analysis), spatial autocorrelation, slope, aspect and curvature calculation, automatic drainage pattern derivation and many other analytical functions are also developed in the IGIAS to diagnose the quality and error pattern of the DEM grids and to analyze the terrain characteristics for iteratively refining model parameters.

A variety of visualization and mapping functions are an important component in the IGIAS. Vector dominated topographical source data, raster-based DEM product and

remote sensing images can be simultaneously displayed to perform graphical overlay analysis. The intermediate results and final DEM grids can be readily visualized and displayed as hill-shaded images, synthetic stereo images, perspective wire-frames, gray scale or color filled contour bands, and contour lines. In conjunction with satellite images or other raster layers, they can be also rendered as color images using a RGB or HSV color model. Zoom-in, zoom-out and roaming functions allow for inspecting the data at different scales and locations. In addition, scatterplots, profiles, histograms and other graphs allow visual appraisal of topographical source data and DEM grids. Visualization results can be used as feedback information for a quick assessment about data quality, model selection and parameter specification.

A separate software package SpaceView is also developed using the X-window library, Motif and C programming language (Figure 2.3). It is intended to extract high-resolution elevation measurements directly from satellite image data for updating and enhancing the DEM. Currently, SpaceView includes functions for image radiometric manipulations, least squares surface fitting, three image matching algorithms, and two shape-from-shading algorithms. In addition, the Alaska SAR Facility (ASF) InSAR software (Guritz, 1998) is used to perform interferometric SAR analysis over the Dry Valley. The Vexcel RaST Radar Stereo Toolkit (Marra et al., 1998) is used to extract elevation data over Terra Nova Bay area. The processing results from SpaceView and these two stand-alone software packages are linked with the IGIAS through internal data conversion routines.

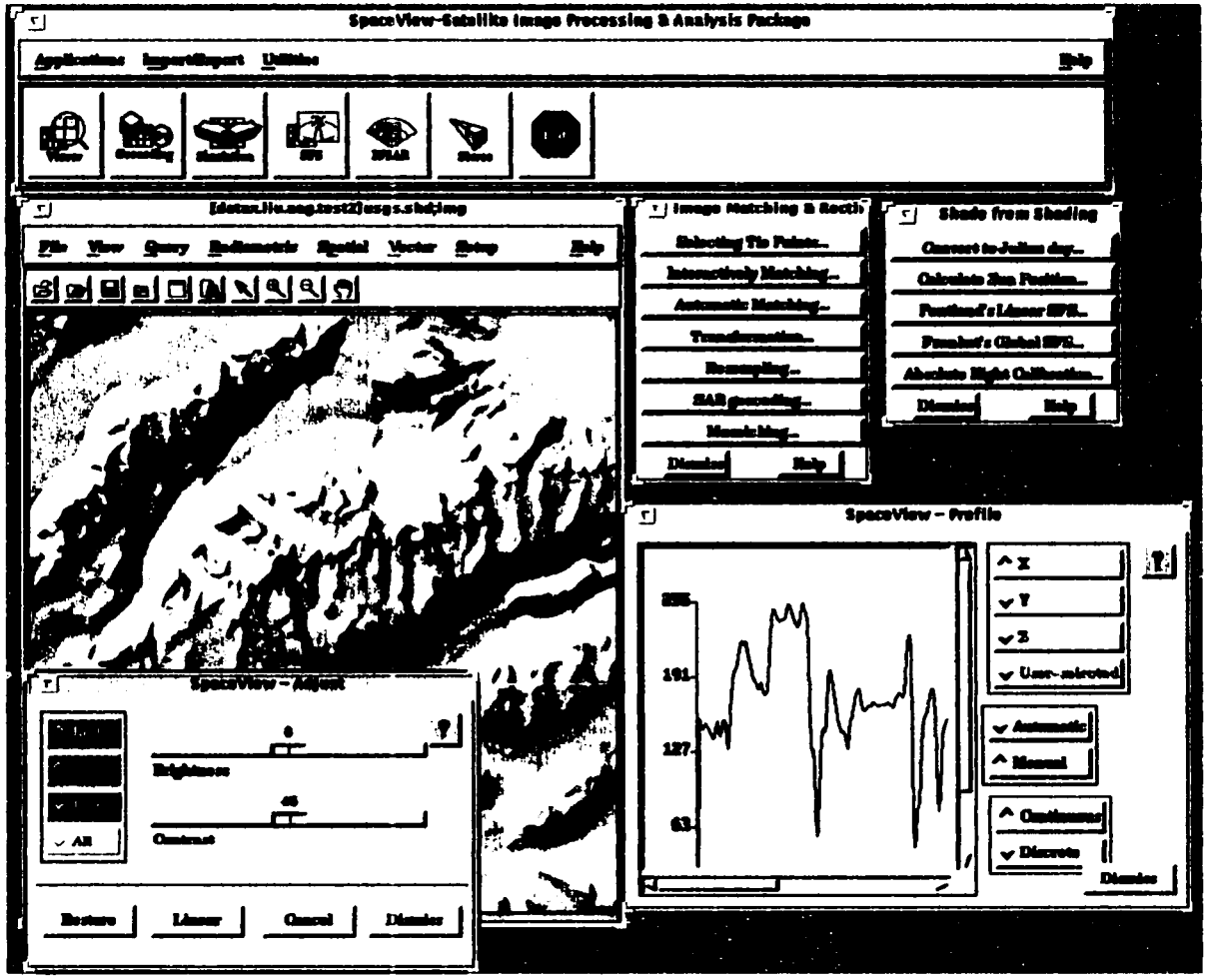


Fig. 2.3: Graphical User's Interface (GUI) for the SpaceView software

As will be demonstrated in the following chapters, the GIS-based software package IGIAS in conjunction with the SpaceView, ASF STEP tool and Vexcel SAR Stereo Toolkit provide a satisfactory approach to processing, integrating and modeling various topographical source data. This integrated system has greatly automated our continent scale DEM generation process.

CHAPTER 3

CONSISTENT AND ERROR-FREE TOPOGRAPHICAL DATABASE

3.1 Topographical database

Elevation measurements captured by various techniques may be recorded and stored in different ways. The essential prerequisites for integrating diverse topographical information are to identify available topographic data sources and to construct a consistent and error-free database. Topographical source data entered into the database need to have a compatible data structure and format so that they can be stored, retrieved and manipulated in a unified GIS environment. It is also desirable to adopt a common horizontal coordinate system and vertical datum within the topographical database. This ensures that different data sets can be spatially correlated and that height measurements from different data sets are directly comparable. The removal of errors and inconsistencies in the topographical database is a critical step for preventing error propagation and generating a reliable DEM product out of diverse sources.

3.2 Characteristics of different data acquisition techniques

An appropriate use of topographical data sets needs the knowledge of corresponding data acquisition techniques. At present a wide spectrum of techniques available for acquiring

digital elevation data. The ground survey and aerial photogrammetry methods represent the traditional elevation extraction and measurement techniques. Contour-based topographical maps have long been the primary storage media for terrain information throughout the world. Since the 1980s, the development of digital photogrammetry workstations and diverse image matching algorithms has greatly automated the conventional comparator-based or analytical plotter-based stereo photogrammetric technique (Miller et al., 1992; Dowman et al., 1992; Leberl, 1990; Lemmens, 1988; Claus, 1984; Heipke, 1992). Meanwhile, the principles and methods of aerial photogrammetry have been extended to satellite stereoscopic images, both optical, most notably SPOT (Day and Muller, 1988; Giles and Franklin, 1996; Chen and Rao, 1993), and Synthetic Aperture Radar (SAR) (Leberl et al, 1986; Leberl 1990; Prati and Rocca, 1990; Dowman, 1991). Also, the shape-from-shading technique has been applied to single airborne and spaceborne images for deriving elevation measurements in combination with some boundary constraints (Horn and Brooks, 1986; Frankot and Chellappa, 1987; Thomas et al., 1991; Guindon, 1990; Cooper, 1994; Fiksel, 1994). The newly developed interferometric SAR (InSAR) technique has demonstrated great promise for extracting highly accurate elevation data and detecting very small surface displacements (Zebker et al., 1994; Zebker and Goldstein, 1986; Mrstik et al., 1996; Kooij, 1996). Spaceborne radar altimeters have been deployed onboard GEOSAT, SEASAT and ERS-1 platforms for acquiring surface height measurements (Zwally et al., 1983; Ridley et al., 1993; Partington et al., 1991; McIntyre, 1991). Airborne laser altimeters have been implemented and placed into use. Spaceborne laser altimeter is currently under development (GLAS Science Team, 1997; Schutz, 1998). In addition,

the Global Positioning System (GPS) technology has been widely used to accurately measure three-dimensional positions of various objects based on the constellation of GPS satellites. The advent of new data capture technologies has increased the acquisition speed of elevation data, improved the position and height measurements, extended their ground coverage, and reduced the cost.

In terms of data acquisition techniques, topographical data sources can be classified into three broad groups:

- Survey data: including ground-based leveling and satellite-based GPS data;
- Remotely sensed data: including radar altimeter data, laser altimeter data, optical and SAR stereoscopic image pairs, complex interferometric SAR data, and shade information in single images; and
- Cartographic data: including contours, spot height points and surface structural lines digitized from paper topographical map sheets.

In comparison, each data source listed above has its own inherent strengths and limitations in terms of the availability, accuracy, sampling density and pattern, and the ground coverage. Therefore, they might assume different roles in the creation of a large scale DEM. Ground survey and satellite-based GPS technology tend to generate very accurate positional and height measurements. With the Precise Positioning Service (PPS), the horizontal and vertical accuracy of GPS measurements are better than 16 m and 23 m, respectively. The differential GPS (DGPS) technique that utilizes two or more GPS receivers simultaneously can achieve the centimeter level of horizontal and vertical accuracy for a 10 km baseline using a double or triple differencing algorithm (Hofmann-Wellenhof et al., 1994; Parkinson et al., 1996). Since both ground survey and GPS

techniques require a physical visit of ground sampling points, the resulting measurements are usually sparse. Often, they are utilized as GCPs (Ground Control Points) for extracting more dense elevation data from stereo photogrammetric and InSAR techniques or used as checking points for the DEM accuracy assessment.

Radar altimeters can provide accurate elevation measurements regardless of weather conditions over ocean and relatively flat and low-slope ground surface, but are prone to errors over highly sloped and rugged lands due to the relatively large footprint (Martin et al., 1983; Brenner et al., 1983). The sampling points of radar altimeter are densely distributed along the flight traverses, but widely spaced between traverses. A laser altimeter can provide height measurements of submeter level accuracy with the aid of GPS and Inertial Navigation System (INS) in determining the position and attitude of aircraft or spacecraft. Compared with radar altimeters, laser altimeters have much smaller footprints and negligible penetration depth, and therefore provide more precise height measurements even on a sloping and undulating surface (Csatho et al., 1996; GLAS Science Team, 1997). Due to the relatively large across-track separation, the spatial density of laser altimetry measurements is limited. In addition, the performances of laser altimeters are influenced by weather conditions.

Stereo photogrammetry and InSAR techniques can be applied to both aerial photographs and satellite image data to extract dense and accurate elevation measurements. In general, the airborne technique results in better resolution and higher accuracy, while the spaceborne technique can achieve much larger ground coverage in a relatively short span of time. Both stereo and InSAR techniques require high quality GCPs to derive absolute position and height measurements. The performance of the

stereo technique depends on many factors, including base-height ratio, the flying height, and image texture (Ehlers et al., 1989; Vincent et al., 1987; Bolstad et al., 1994; Leberl 1990). The InSAR technique is mainly influenced by baseline length, image signal/noise ratio, temporal decorrelation (for repeated-pass InSAR), and geometric distortions (layover, shadow, and excessive foreshortening) (Lin et al., 1994; Zebker et al., 1994; Goldstein et al., 1989). The shape from shading method can derive pixel level high-resolution height measurement based on image shading information and boundary constraints. Nevertheless, most shape from shading algorithms assume that the terrain surface is continuous and has a uniform albedo (Frankot and Chellappa, 1989; Horn, 1990; Cooper, 1994). Consequently, they are better suited for fairly uniform and smooth land cover, such as ice and snow covered polar regions and sand covered deserts.

Cartographic data are a second hand topographical information source, in the sense that they are originally derived from direct height measurements of some kind as described above. During the map making process, the small-scale features contained in the original measurements are often filtered out, resulting in elevation data of a degraded quality. The information content of a topographical map depends on not only the quality of the original source data, but also map scale and contour interval. In addition to contours and spot heights, topographical maps may also contain, implicitly or explicitly, many important terrain features such as surface break lines, ridges and drainage lines. In the long history of ground-based survey and stereoscopic measurements of aerial photographs, a large volume of topographic maps have been created at various scales throughout the world. Due to the relatively low cost and the widest availability, the

digitization of topographical maps represents the practical method for gathering digital elevation data for a large area, particularly for national and continental scale projects.

3.3 Topographical data sources for Antarctica

As a case study, a comprehensive topographical database was compiled for creating a continental scale DEM over the Antarctic. It contains cartographic data from the Antarctic Digital Database (ADD), the USGS and the Australian Antarctic Division, satellite radar altimeter data, and airborne radar data. In support of DEM quality verification and other relevant operations, various satellite image data are also collected and incorporated into the database, including the USGS AVHRR mosaic (Ferrigno et al., 1996) and the Radarsat SAR Quicklook mosaic (Jezek, 1998) over the entire continent, DISP photographs, and Landsat and SPOT images at selected locations.

The ADD represents the best digital collection of the existing cartographic data over the Antarctica. It was compiled by the British Antarctic Survey (BAS) in collaboration with the Scott Polar Research Institute (SPRI) and the World Conservation Monitoring Center (WCMC) (Cooper, 1993; BAS, SPRI and WCMC, 1993). Among 210 map sheets digitized, 164 are at the scale of 1:250,000. As shown in Fig. 3.1, the cartographic data cover the entire continent. In general, the scale and density of cartographic data are larger over the mountainous and coastal regions, especially over the Transantarctic Mountains, the Antarctic Peninsula, the Ellsworth Mountains, the marginal coasts of the Marie Byrd Land and the Queen Maud Land. In the interior of Antarctica, cartographic data have a smaller scale and sparse contour coverage, hence showing much less topographical details. Among 18 different thematic coverages, those containing contour, spot elevation

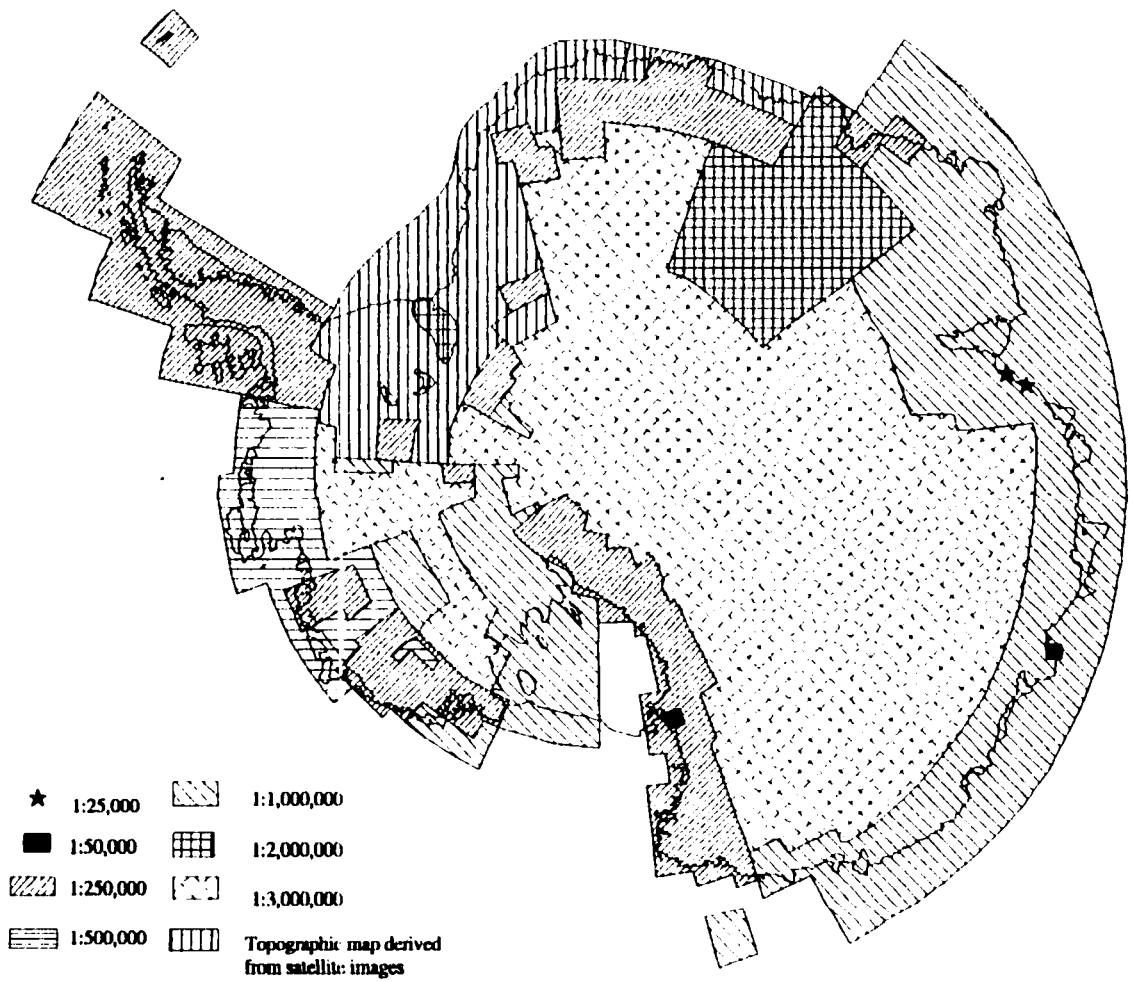


Fig. 3.1 : Scale and coverage of cartographic data

points and coastlines are used as direct input data, while those containing the grounding lines, streamlines, and lake shorelines are used in our modeling process to constrain the spatial interpolation for producing topologically and geomorphologically correct DEMs. The data coverages are available at five different scales. The most detailed data, namely, the source topographical maps of the original scale are used.

To augment the ADD, a considerable number of topographic maps at larger scales were acquired from the USGS, the Australian Antarctic Division, and Institut für Angewandte Geodäsie of Germany. Thirteen topographical map sheets were manually digitized by the USGS at a scale of 1:50,000 with a contour interval of 50 m (Hallam and Mullins, personal communication), covering the Dry Valley region of the Transantarctic Mountains. The Australian digital contour data were captured from a number of large-scale topographical maps (Belbin and Ryan, personal communication). Three digital contour data sets that we used cover the Vestfold Hills, the Larsmann Hills and the Windmill Islands in the East Antarctica at the scales of 1:25,000~1:50,000 with a contour interval of 10 m. We also digitized the contour lines over Berkner Island, Henry Ice Rise and Korff Ice Rise in the Filchner-Ronne Ice shelf from a 1:2,000,000 German topographical map.

Three sets of elevation data derived from ERS-1 radar altimetry over the Antarctic were acquired from different investigators. A radar altimeter is a nadir-pointing active microwave sensor designed to measure the surface height over ocean and ice surfaces. The radar altimeter transmits a short duration Ku-band pulse vertically downwards and then tracks the returned radar pulse. The information on the shape and timing of the returned signal is utilized to estimate the surface height. The radar response over the

continental ice sheets is considerably more complex than over the oceans, as the typical return echo has an unpredictable shape and the land topography has a larger variability than the ocean surface. To obtain valid elevations for ice sheets from satellite altimeters, a series of corrections are needed, especially, retracking of the timing error and slope-induced error correction (Martin et al., 1983; McIntyre, 1991; Brenner et al, 1983). The spatial distribution of the radar altimeter measurements is highly anisotropic. Height sampling points are separated by 335 m along-track and by up to 4 km across-track at the latitude of 60°S, reducing to 2 km at 70°S (Bamber and Bindschadler, 1997). Satellite altimeters have a footprint (3-dB beam-limited) size of up to 21 km, and the resulting point measurement after correction represents the average height within the first pulse-limited footprint (Zwally et al., 1997). The wide across-track separation and large footprint lead to the relatively coarse spatial resolution.

The orbital inclination of ERS-1 is 98.6° and its corresponding data coverage in the Antarctic is limited to the region north of 81.4°S. Ihde et al. (1995) used 17 full 35-day repeat cycle data sets and generated an elevation data set with a latitude spacing of 0.05° and longitude spacing of 0.2° using the bicubic spline interpolation method. Bamber applied the offset center of gravity method for wave-form retracking and the relocation method for slope-induced error correction to two 168-day cycles of the ERS-1 geodetic mission, and created a 5-km spacing elevation grid using the distance-weighted and TIN interpolation methods (Bamber and Bindschadler, 1997). Zwally et al. (1997) used a 9 or 5-parameter fitting function for waveform retracking and a direct slope correction method for slope-induced error correction. Based on two 168-day cycles from the geodetic mission and two 35-day cycles from the multi-disciplinary mission, an elevation data set

of approximately 5 km spacing was generated by using a bi-quadratic function fitting algorithm. By combining Geosat and Seasat altimeter data, Lingle et al. (1994) produced a 3 km resolution elevation grid over the Amery Ice Shelf in East Antarctica using the Kriging algorithm.

To obtain the mean sea surface height over the off-shore ocean area with reference to the WGS84 ellipsoid, we also acquired the ERS-1 altimeter ocean level 2 product (OPR) from the ESA (European Space Agency), which were derived from a 168-day cycles of the geodetic mission. The accuracy is estimated to be significantly better than 10 cm (Martin et al, 1983).

During the 1988-89 field season of the Siple Coast Project, over 14,000 km of high-quality digital airborne radar sounding data were collected over the upstream parts of Ice Streams A, B, and C in the West Antarctica (Retzlaff, Lord and Bentley, 1993), where there is no satellite altimeter data available. The flight lines were regularly spaced 5-10 km apart, and the data were densely sampled along the flight lines with an along-track spacing of roughly 120 m. The radar system was a modified SPRI Mark IV 50 MHz radar mounted in a DeHavilland Twin Otter aircraft equipped with a Litton LTN92 inertial navigation system (INS). The error estimates from cross-over analysis and ground station validation is about 4~9 m.

Airborne and station radar sounding ice thickness data were collected over the Ross Ice Shelf during the RIGGS (the Ross Ice Shelf Geophysical and Glaciological Survey) survey started in 1973 (Bentley and Jezek, 1981). A series of radar echo sounding (RES) flights by the Twin Otter aircraft were conducted from each of the base camps, which were arranged roughly on a 55 km grid. RIGGS ice thickness data were digitized from

an ice thickness map (Bentley and Jezek, 1981), which have an accuracy of about 10 m (Greischar et al., 1992). Ice thickness data were converted to surface elevation by a linear hydrostatic equilibrium model.

Airborne radio echo sounding data were captured by the British Antarctic Survey (BAS), covering the area around the Evans Ice Stream and Fowler Peninsula. The aircraft was navigated by differential carrier phase GPS. The cross-over analysis gives an average error of 6 m with a standard deviation of 7 m (Jones, personal communication).

Antarctica has relatively poor elevation coverage by conventional ground surveys. The available ground survey data are mainly contained in the supplemental spot height point layers in the ADD (Fig. 3.2). The spot elevation points are sparsely distributed over space and contain various types of height measurements, including differential leveling, trigonometrical, airborne altimetric, surface barometric, photogrammetric, geodetic satellite observation station, astronomical station, and survey control station. GPS measurements that we assembled are several traverses over the Siple Dome (Scambos, personal communication), the Lambert Glacier Basin and the Amery Ice Shelf (Kiernan, 1998), the Rutford Ice Stream (Vaughan, personal communication), and scattered points in the Transantarctic Mountains and the West Antarctica (Whillans, Berkman and Wilson, personal communication) (Fig. 3.2). Ground survey and GPS data are spatially limited and not dense enough to create a grid DEM alone. Therefore, they are incorporated with other data in the DEM computation, or reserved as ground truth for the final DEM accuracy evaluation.

Using the IGIAS, the collected source data were converted into topologically structured ARC/INFO coverages. Contours, coastline and grounding lines are represented

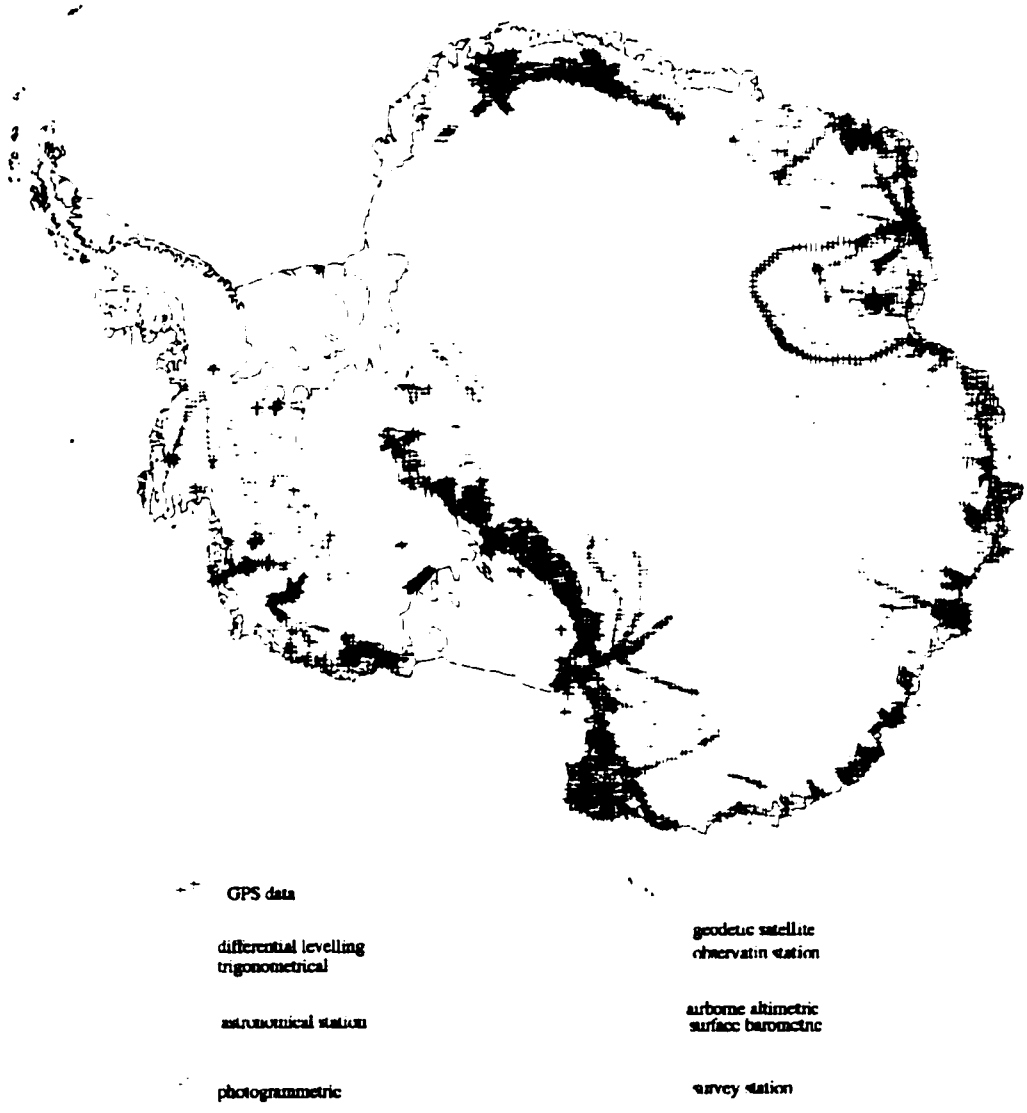


Fig. 3.2: Distribution of ground survey and GPS data

as ARC/INFO line coverages, while the spot height points, satellite and airborne radar altimeter data, and GPS data are stored as ARC/INFO point coverages. Height values and other attributes are contained in the Arc Attribute Tables (AAT) or Point Attribute Tables (PAT) associated with the appropriate coverages. Satellite image data are converted into grid layers. The coherent data structure and tight coupling between spatial entities and their attributes enables a variety of GIS operations that are required in the DEM generation process.

3.4 Planimetric and vertical reference systems

The location of elevation measurements can be defined in the true spherical geographic coordinate system in terms of latitude and longitude (φ, λ) or by two dimensional plane coordinates of a map projection (x, y). Latitude and longitude are measured with the Prime Meridian and the Equator as the reference planes based on an earth ellipsoidal model. Map projections portray a portion of the earth on a flat surface. The transformation from 3D space to 2D involved in map projections inevitably distorts at least one, often more, geometric properties, such as shape, area, distance, and direction of terrain features.

In the case of Antarctica, the Polar Stereographic projection with a standard latitude of 71°S and a central meridian of 0° is selected as the common planimetric reference system. The coordinates are in meters with the origin at the South Pole. With this conformal projection, the shapes of topographic features can be correctly represented, but areas and lengths measured at different latitudes are not directly comparable. The topographic data sources that we collected for Antarctica were referenced to various

planimetric coordinate systems, including the true spherical geographic coordinate system, polar stereographic projection, UTM, and Lambert Conformal Conic projection. Using the IGIAS project functions, all the source data are transformed into the coordinates of the Polar Stereographic map projection.

A vertical datum is the zero surface to which height measurements are referenced. Mean Sea Level (MSL), geoid and ellipsoid are often used as vertical datums, and the corresponding height measurements are respectively known as elevation, orthometric height and ellipsoidal height. MSL is the average surface of the oceans, and it is commonly used as the vertical reference surface in conventional surveying and mapping to measure the variation of topography. The geoid is the equipotential (level) surface defined by Earth's gravity. Local variations in the earth's core and surface materials cause the gravity surface to be irregular. Orthometric height is measured along the plumb line from the ground point to the geoid (Hoar, 1982; Schreier, 1993). Since it is very difficult to accurately measure the irregular gravity shape, the geoid has been roughly approximated in the past by the MSL. The difference between the geoid orthometric height and MSL elevation, which is caused by ocean currents and tidal effects (dynamic ocean height), is normally less than 2 m (Smith, 1996). In Antarctica, the MSL is approximately 1.5 m below the geoid due to the circumpolar currents of the Southern Ocean (Bamber and Bentley, 1994; Bamber and Bindschadler, 1997; Rapp, 1991).

With the development of satellite surveying, especially GPS and satellite radar altimetry techniques, the ellipsoidal height is now used in many cases to represent the terrain vertical height. The ellipsoidal height (geodetic height) is the vertical distance of the point on the terrain surface above or below the reference ellipsoid (Hoar, 1982;

Schreier, 1993). The ellipsoid is the geometric model of the earth, which is centered on the earth's center of mass and defined by two parameters. One is the semi-major axis dimension and the other parameter is the semi-minor axis dimension, flattening, or eccentricity. The difference between orthometric height (H) and ellipsoidal height (h) is the geoidal undulation (N), namely, the height of the geoid above the ellipsoid (Fig. 3.3). The mathematical relationship is as follows:

$$h = H + N \tag{3.1}$$

In the Antarctic, the geoidal undulation N ranges from –67 m to +42 m, and its spatial pattern is shown in Fig. 3.4. It was derived from the OSU91A geoid model over the continent and approximated by the Mean Sea Height derived from the ERS-1 radar altimeter ocean mode data over off-shore ocean areas.

Among the topographical sources collected for the Antarctic, cartographic data are referenced to the Mean Sea Level (MSL), while the GPS data, satellite radar altimeter data and airborne radar data are referenced to the WGS84 (World Geodetic System 1984). In addition, GRS80, WGS72 and Rapp Geoid A are also used in some source data as vertical datums. In datum adjustment, we ignored the differences between the MSL elevation and the geoid orthometric height. The most recent geoid model OSU91A (Rapp et al., 1991) is used as the common vertical datum to integrate all the source data. This is because the WGS84 ellipsoidal heights of the radar altimeter data and GPS data are point type measurements and easier to convert to the orthometric height relative to the OSU91A geoid, but the contour-based cartographic data are relative to the MSL and difficult to be directly converted to the WGS84 ellipsoid heights due to the variation of geoidal undulations along contour lines.

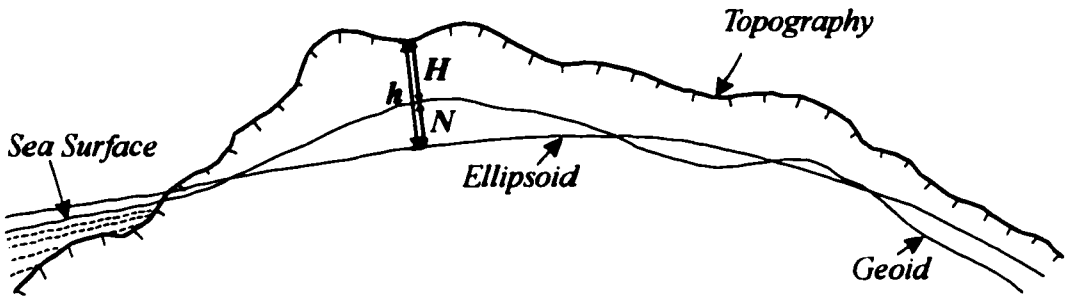


Fig. 3.3: Relation of topography, geoid and ellipsoid

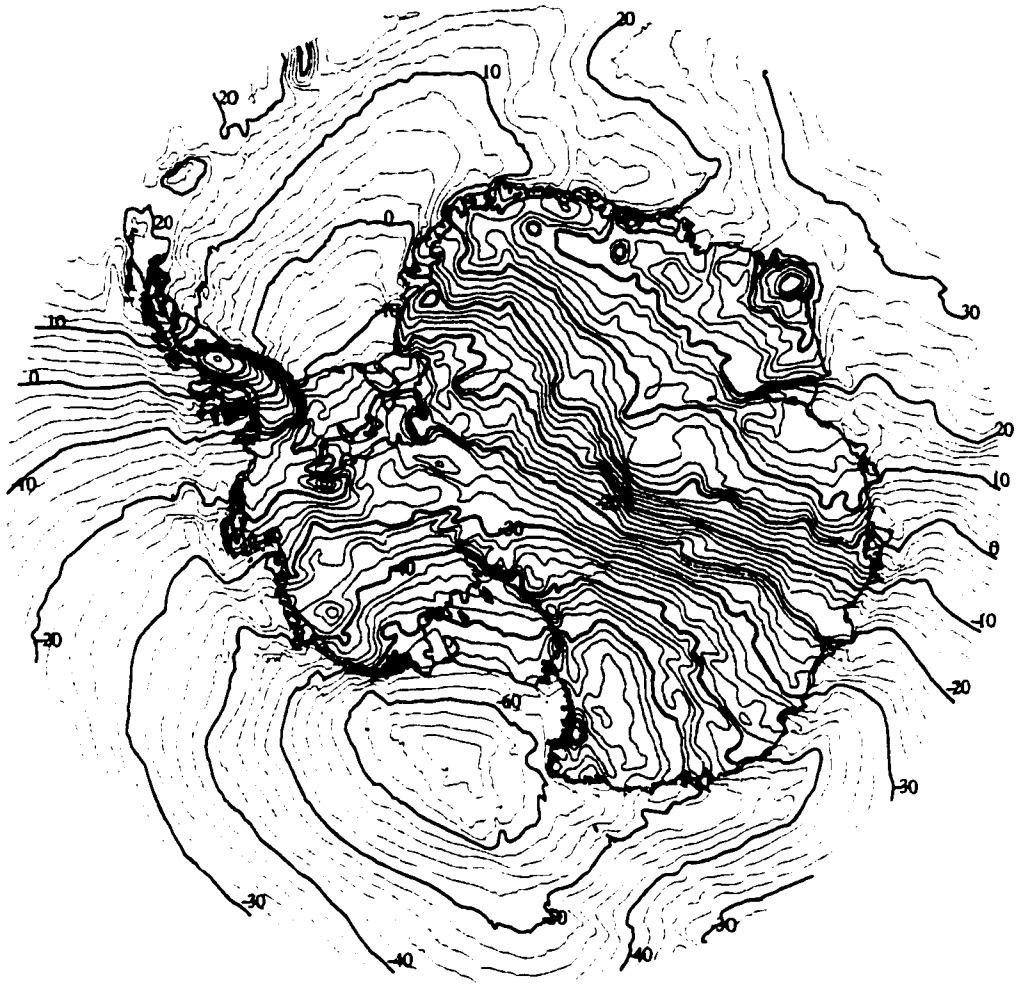


Fig. 3.4: Geoid undulation of the Antarctic. Contour interval is 2 m.

For the final product, we first computed the orthometric height DEM grid referenced to the OSU91A geoid by interpolating the unified source data. Then, the geoidal undulation grid is added to the orthometric DEM, resulting in a new ellipsoidal height DEM relative to the WGS84. The orthometric height DEM is often used in geophysical applications due to its physical significance, while the ellipsoidal height DEM is most convenient for remote sensing applications.

3.5 Error detection and correction techniques

3.5.1 Errors in topographical data

Errors may arise during data acquisition and subsequent data handling and processing. The possible causes include malfunction or inadequacies of the measuring instrument, a mistaken reading or recording, calculation and execution faults, digitizing errors, and so on. The detection and elimination of erroneous values in source data is obviously important for maintaining a high quality topographical database and preventing error propagation in the DEM generation process.

We examined all data layers in the Antarctic topographical database and detected various errors. In contour coverages, there exist two primary types of errors: mislabeled contours and intersections of contours. A contour is digitized as a sequence of line segments, and each segment of the contour line was assigned an elevation value and recorded as an independent entity. The errors of assigning wrong elevation labels lead to the situation where different segments of the same contour line have different elevation values or the whole contour has an inconsistent elevation value with its neighboring contour lines. In some cases, the digitized contour lines deviate from their original

position on the source map and intersect each other, resulting in some positions having two or more conflicting values. In addition, the poor ground control and inaccurate navigation techniques used to acquire the original topographical data can also cause positional errors of contours.

In the spot height coverages, some geodetic control points, which only have the measurements on horizontal positions, were simply assigned zero elevation values during the digitizing process (Cooper, personal communication). Some points were mistakenly assigned extremely large values out of the reasonable range due to data entry errors. The satellite radar altimeter data were preprocessed by other investigators. Although retracking, slope correction and other filtering procedures have been applied (Bamber and Bindschadler, 1997; Zwally et al., 1997), some errors remain in the data sets. The remaining errors may be caused by poor tracking or complete loss of the returned echo in the original measurements, and may be induced by the imperfections of the slope correction and spatial interpolation algorithms. In the airborne radar echo sounding data, erroneous elevation values were also found.

Due to the large volume of source data, manual inspection and editing would be tedious and prohibitively time-consuming. In this research, a number of interactive and automatic techniques were developed for detecting and correcting anomalous values and gross errors in source data.

3.5.2 Global statistical analysis and cross-validation methods

Global statistical analysis is based on summary statistics of all the measurements in one topographical data set. In our Antarctic topographical database, point type measurements

are represented as ARC/INFO point coverages, and height values and other attributes are stored in the associated Point Attribute Tables (PATs). Contours are represented as ARC/INFO line coverages, and elevation labels and other attributes are stored in the associated Arc Attribute Tables (AATs). A variety of summary statistics can be derived from the PATs and AATs, such as the mean, standard deviation, median, maximum and minimum values, and frequency distribution (histogram). These global statistics are very helpful in identifying extreme erroneous values. For a specific region, we can infer reasonable minimum and maximum elevation values, namely the elevation range, from prior knowledge about this region or from the frequency distribution of elevation measurements. The data points that have elevation values out of the reasonable range can be selected and flagged using logical operations on the PATs or AATs. Using this method, we removed negative elevation values (relative to OSU91A Geoid) and extremely large values in satellite and airborne radar data coverages and spot height coverages. In spot height coverages, incorrect zero elevation values over the continent that are away from the coastlines are considered wrong and discarded.

For contour coverages, all possible and regular elevation values can be inferred from the elevation of the base contour line and contour interval. Using the AATs, we easily identified contour lines that had spurious and irregular elevation values. They are corrected to right values inferred from neighboring contours.

When multiple topographical data sets are available for the same area, cross-validation analysis can be conducted to detect errors and enhance the consistency and reliability of the topographical database. In the case of Antarctica, spot height coverages and contour coverages spatially overlap each other. We developed an automatic method

by using the ARC/INFO internal functions to check spot height points with reference to the corresponding contour coverages. This method first predicts the elevation values at the positions of spot points by interpolating the contours, then computes the differences between the interpolated values from contours and elevation values recorded in spot height coverages. This is computationally realized through an ARC Macro Language (AML) program to combine interpolation, "latticespot", and grid (map) algebraic operations. We removed those points that have an absolute difference greater than one contour interval in the flat area and two times greater than the contour interval in the rugged area. Similarly, we also performed the cross-checking between contour data and satellite radar altimeter data in the GIS environment.

The comparison between different versions of satellite radar altimeter data revealed systematic errors in an earlier version provided by one investigator due to their confusion of the vertical reference system. The magnitude and distribution of the difference between satellite radar altimeter data from different investigators also provide an indication of the reliability of satellite altimeter data and the dependency of discrepancies between different processing techniques on the terrain characteristics. For example, we compared the ERS-1 data pre-processed by Zwally et al. (1997) with that by Bamber (Bamber and Bindschadler, 1997). We observed that the discrepancy is about 1.5 m over the ice shelves. Most regions show agreement to better than 20 m, but significant differences in excess of 50 m are observed between the two data sets in marginal areas where the surface is relatively rough and steep.

3.5.3 Scientific visualization and image simulation methods

Visual inspection and interactive editing of topographical data can be conducted in

various ways. The contours and point type measurements can be displayed using different symbols in terms of their elevation values or other attributes in the PATs and AATs. To detect the mislabeled and intersected contours, we developed a function in the IGIAS that allows for displaying a number of adjacent contour lines by a specified color sequence. A line segment that has a different color from the rest of the contour line or intersects with contour lines of other colors indicates a wrong label or intersection error.

Errors in the source data can be also detected and traced by visually inspecting the interpolated DEM grid or its derivatives by means of a variety of rendering techniques, such as a perspective view, contour lines, hill shading, and synthetic stereo display. Perspective displays provide convincing visualization of the terrain surface, and gross errors tend to be projected out of the surface and thus detectable. The drawback of perspective display as an error checking tool is that some parts of the surface are hidden, and it is difficult to inspect the entire surface at a time. Contour lines derived from the interpolated DEM grid can be conveniently overlaid with the source data to uncover the errors. For example, when overlaying the contours derived from the DEM grid with the source contours, the places where the derived and source contours are sharply different often indicate wrong elevation labels in the source contours or inadequacy of the interpolation algorithm.

Hill shading and synthetic stereo displays are particularly useful for revealing errors in source data. In addition, visual comparison and digital image matching between the simulated image and real satellite image can help to evaluate the horizontal resolution of the DEM grid and the geometric fidelity of terrain features in the DEM. This in turn

reflects the topographical information content and positional accuracy of the topographical source data used to make the DEM grid.

To produce a hill-shaded relief image, an image formation model is required to calculate the image intensity *I(x,y)*. The Lambertian model was used. It assumes that the terrain surface has a uniform albedo and diffuses the solar illumination equally in all directions. Under this assumption, the intensity of diffusely reflected light is proportional to the cosine of the local incidence angle between the surface normal and the illumination vector (Fig. 3.5). The cosine of the local incidence angle equals the dot product of the normalized illumination vector and surface unit normal vector:

$$I(x,y) = \rho \cos i \tag{3.2}$$

$$\cos i = \vec{s} \bullet \vec{n} = \frac{-z_x s_x - z_y s_y + s_z}{\sqrt{z_x^2 + z_y^2 + 1}} \tag{3.3}$$

$$\begin{aligned} s_x &= \cos\theta \cos\varphi \\ s_y &= \cos\theta \sin\varphi \\ s_z &= \sin\theta \end{aligned} \tag{3.4}$$

where i is the local incidence angle between the surface normal and illumination direction vector; ρ is the albedo of the terrain surface; $\vec{s} = (s_x, s_y, s_z)$ is the unit vector of illumination direction; $\vec{n} = (-z_x, -z_y, 1)$ is the surface unit normal vector; z_x, z_y are surface slopes along the x and y direction, calculated from the DEM grid; θ, φ are the specified elevation and azimuth angles of illumination. The intensity value calculated from a DEM grid is in an orthographic projection.

Gross errors in the source data tend to propagate into the DEM grid, leading to the anomalies in surface gradients. When the DEM grid is rendered as a hill shaded image or synthetic stereo image, errors appears as anomalous ditches and scars or erratic peaks or

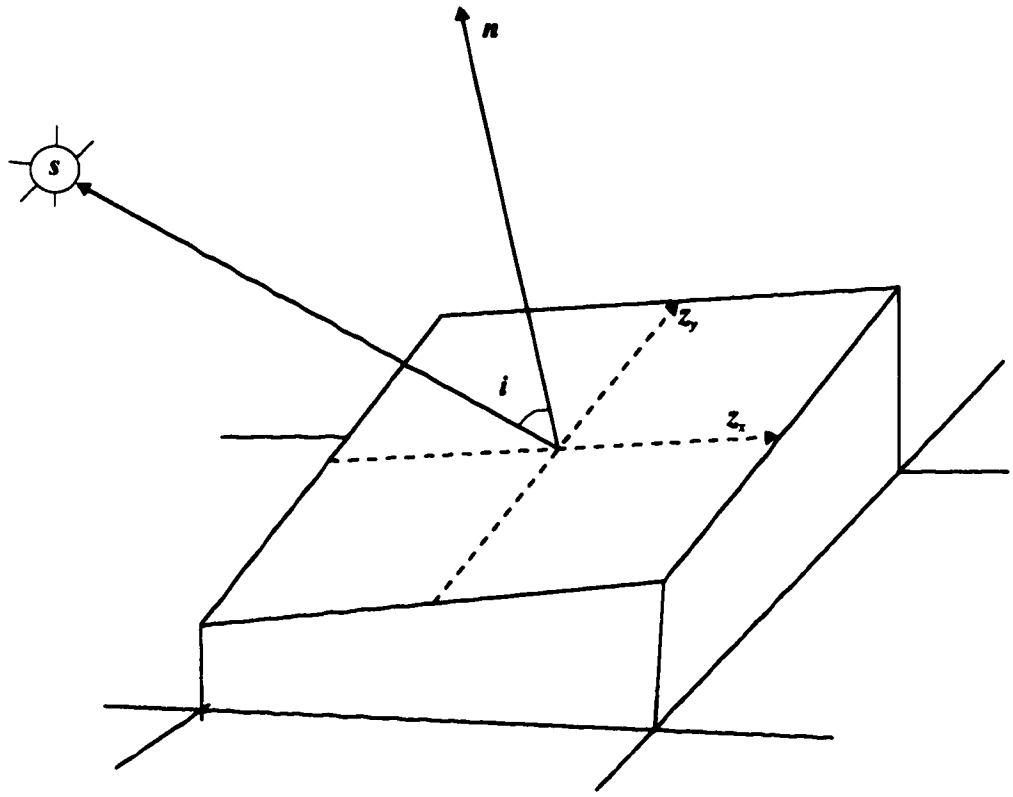


Fig. 3.5: Geometric definition of local incidence angle

pits (Fig. 3.6), destroying the continuity and smoothness of the local terrain surface. Hill shaded images can be rapidly generated from a DEM by using a specified illumination position and vertical exaggeration factor in the orthographic projection. An adjustment of the illumination direction and vertical exaggeration factor can selectively enhance the terrain features with different orientations and sizes. The advantage of an orthographic display of a hill-shaded image is that all parts of the terrain surface are visible and relatively undistorted in comparison with a perspective view.

Based on an arbitrarily hill-shaded image, we can further calculate a synthetic stereo image mate by introducing stereoscopic parallax. The parallax of any position on the terrain surface is directly related to the elevation of that point above the datum (Batson et al., 1976; Jensen, 1980; Usery, 1991):

$$\Delta p = \Delta h(K) \tag{3.5}$$

where Δp is the parallactic displacement of a point on an image whose elevation above the datum is Δh; and K is a constant that determines the strength of the stereoscopic illusion of depth. Large values of K result in more vertical exaggeration of terrain. With the simulated image and its synthetic stereo image mate, an anaglyph stereo image can be composed using a RGB color model and viewed through a pair of red-blue glasses with improved capability in the perception and identification of surface anomalies. An example is shown in Fig. 3.6.

During the Antarctic DEM generation process, all the intermediate individual DEM grids have been hillshaded. Then, topographical source data, either contour-based cartographic data or point type of radar altimeter data, are overlaid on the top of the hill-shaded gray scale image or the anaglyph stereo image. Using this method, numerous

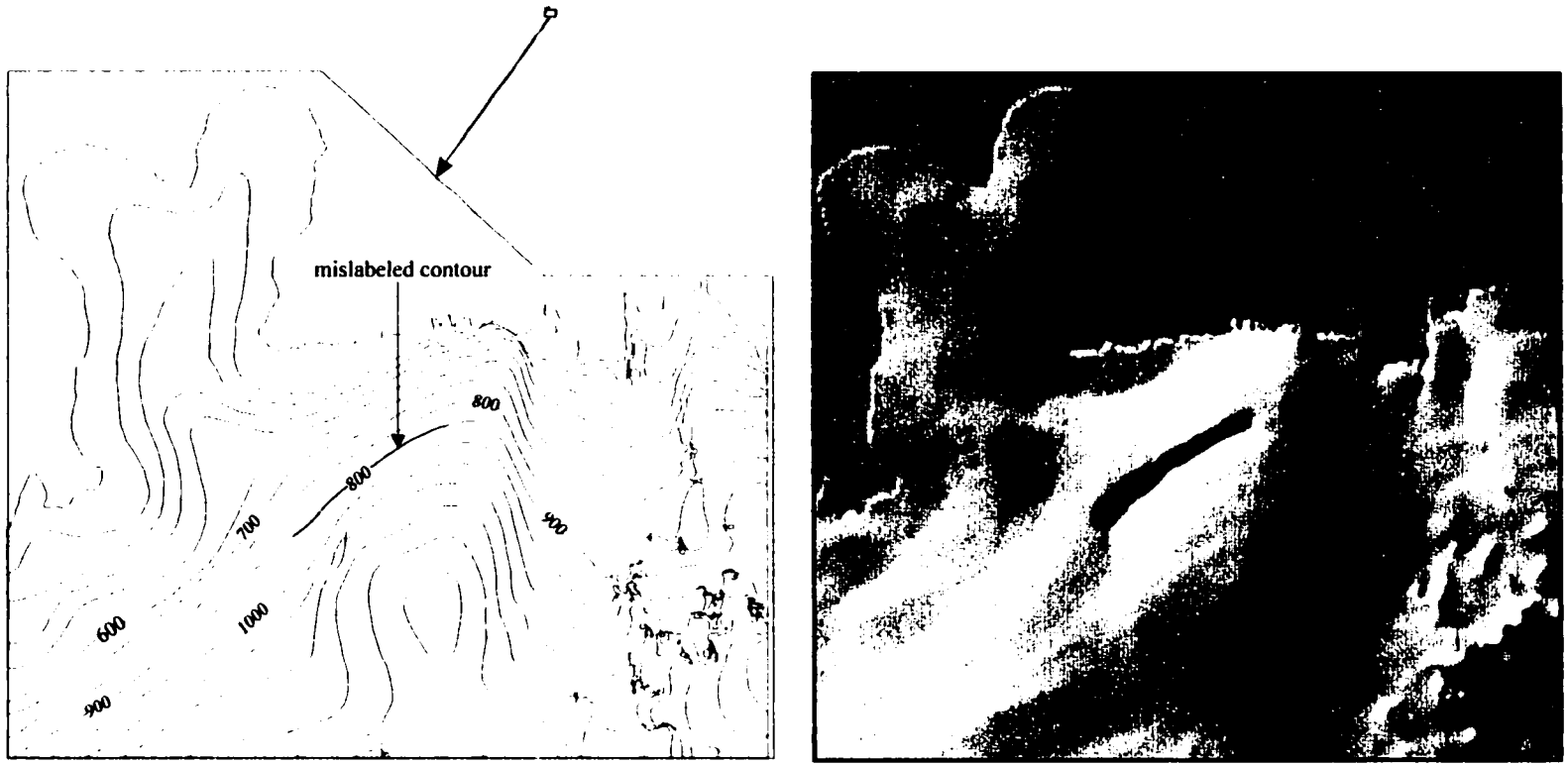


Fig. 3.6: A mislabeled contour detected in hill-shaded image through synthetic stereo view. (a) Original contours; (b) Synthetic stereo image. Stereo view can be perceived with a pair of red-blue anaglyph glasses

wrong labels and intersections in contours and random gross errors in point type radar data have been detected as they appeared on the simulated gray scale image or stereo image as unnatural abrupt artifacts.

When real optical or SAR images exist, it is possible to perform a rigorous image simulation. Image simulation refers to the digital synthesis of a satellite image according to a DEM grid in conjunction with information about illumination position and imaging geometry contained in the ephemeris data of satellite images. In an area with homogeneous land cover like the Antarctic, the comparison and correlation analysis between simulated image and real image can reveal the errors and imprecision of the DEM in an automated fashion.

To detect and correct positional errors of DEM products, we used a rigorous SAR image simulation technique (Holtzman et al., 1978; Kaupp et al., 1983; Leberl, 1990). According to the radar equation (Ulaby et al., 1982; Curlander and McDonough, 1991), the magnitude of the returned power can be decomposed into two parts, namely, system effect K and terrain effect σ. The terrain effect σ can be further split into two components, the differential scattering cross-section per unit area (backscatter coefficient) σ_0 and the illuminated area of the resolution cell A:

$$P_r = K \cdot \sigma \tag{3.6}$$

$$K = \frac{P_t G^2 \lambda^2}{(4\pi)^3 r^4} \tag{3.7}$$

$$\sigma = A \cdot \sigma_0 \tag{3.8}$$

where P_r is the received power; P_t is the average transmitter power delivered to the antenna during the time of a transmitted pulse; G is the dimensionless two-way gain of

the transmitting and receiving antenna; λ is the transmitted wavelength of the microwave; σ_0 is the differential scattering cross-section per unit area; A is the illuminated area of the resolution cell; and r is the slant range from the radar transmitter to the ground target.

For a specific SAR system, the parameter K can be approximately regarded as a constant. The scattering coefficient σ_0 is a function of local incidence angle and terrain properties, such as roughness, dielectric properties, and moisture content. Given a local incidence angle and specific terrain cover, the scattering coefficient can be taken from an experimentally established backscatter look-up table or determined from analytic reflectivity models. The area effect A can be computed as the cosine of the angle between the local surface normal and the radar illumination vector (Leberl, 1990). In our simulation, a generalized cosine model (Muhleman, 1964; Leberl, 1998) was used to approximate the relationship of returned radar power with the local incidence angle:

$$\sigma = \log \frac{M^3 \cos i}{(\sin i + M \cos i)^3} \tag{3.9}$$

where i is the local incidence angle, and M is a constant coefficient calibrated from real image, 0.1 in our case.

Using the satellite ephemeris data supplied with a SAR image, the instantaneous position (x,y,z) and orientation (velocity vector) of the satellite sensor at the time that a particular pixel was imaged with respect to a geocentric coordinate system were computed. With the SAR imaging geometry and orbital information, two-dimensional arrays of backscattered powers calculated from a backscatter model were projected onto a slant range image. By identifying tie points between the simulated slant range SAR image and

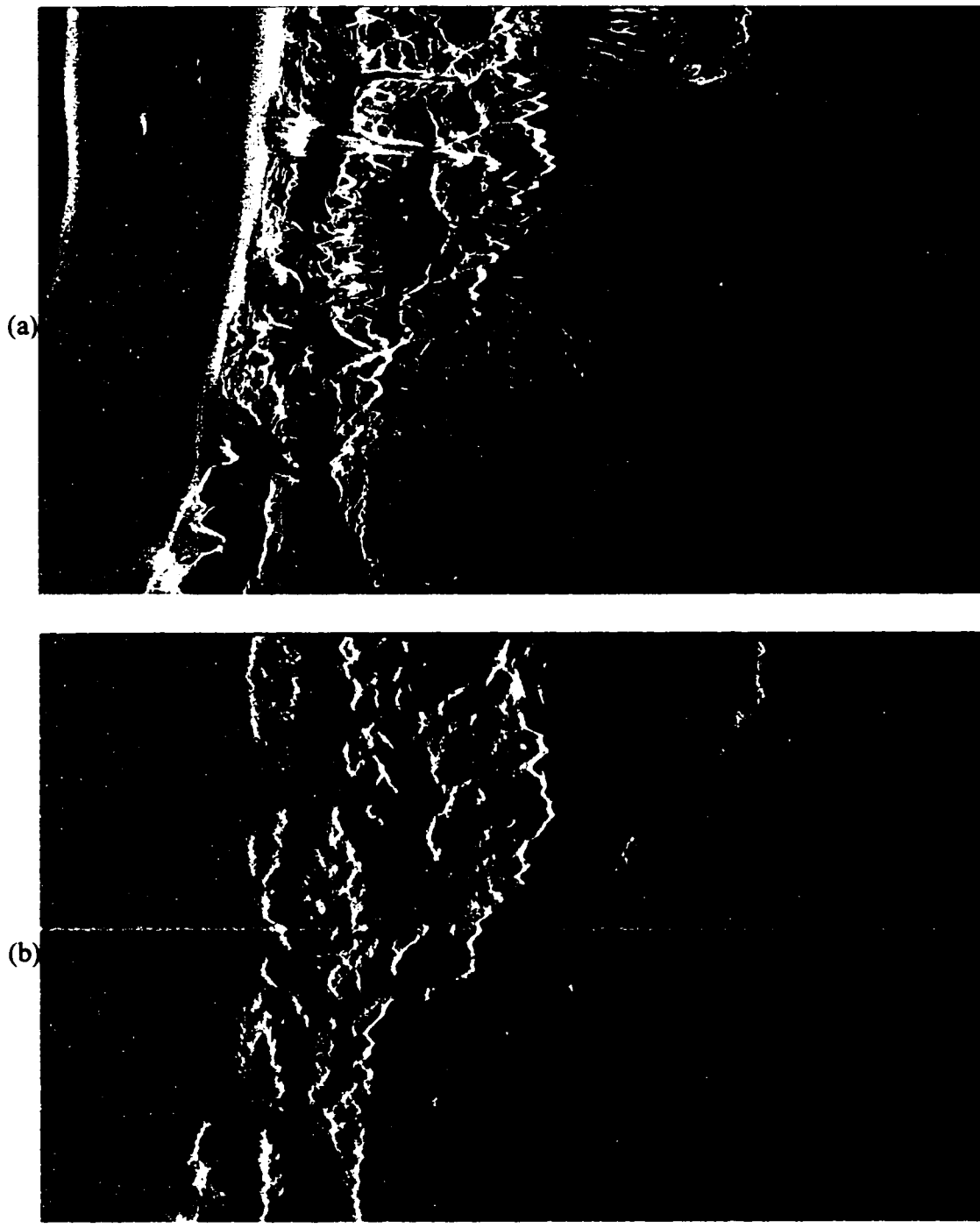


Fig. 3.7: Detection of DEM positional errors using SAR simulation technique.
(a) real SAR image with four distinuishable points A, B, C, D;
(b) simulated SAR image from DEM. The positions of A, B, C, and D calculated from SAR ephermeris data are 3~5km away from the same feature points A′ , B′ , C′ , and D′ .

the real slant range image, we can measure the positional offsets of terrain features present in DEM with reference to the corresponding features in the real satellite image.

A Radarsat SAR image was simulated over the Ellsworth Mountains. Comparison of a simulated image with the real Radarsat SAR image reveals that although the horizontal resolution and relative accuracy of source contours are quite high, the systematic absolute positional error of the source data is significant (Fig. 3.7). The positional offsets of point A, B, C, D in the DEM are as large as 3.8 km, 4.8 km, 4.8 km, and 4.9 km, respectively (Fig. 3.7). The positional errors were caused by the poor ground control and inaccurate navigation techniques during the 1960s when the topographical maps were produced. This positional error was corrected by using a warping equation derived from several selected tie points on the simulated and real satellite images. As Radarsat SAR imagery have been acquired over the entire Antarctic, other rugged areas of the Antarctic will be checked using this technique. The correction of the positional errors of the DEM would increase the coregistration between topographical data and Radarsat SAR images, and hence improve the accuracy of terrain distortion corrections for the RAMP project (Jezek, 1999).

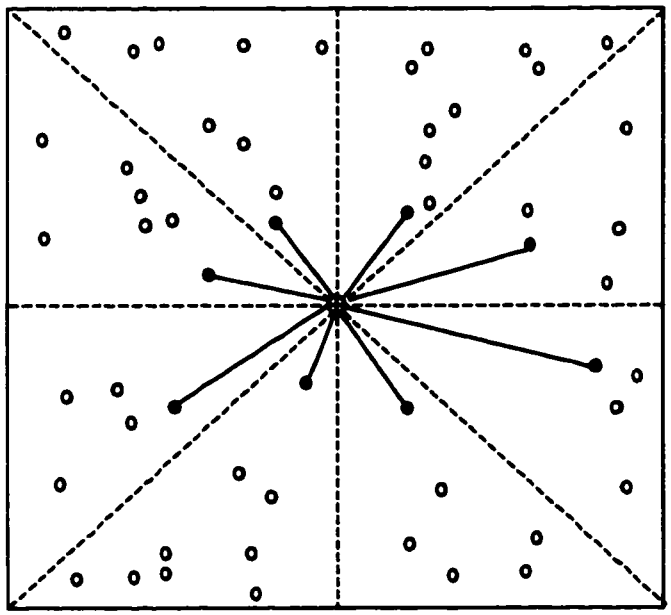
3.5.4 Locally adaptive and robust statistical analysis method

Global statistical analysis is applied to the entire data set. Due to regional heterogeneity of topographical data, a data value that is typical of the whole data set may not be typical within a local neighborhood. Thus, the global method can uncover serious and extreme global errors, but subtle local errors tend to be hidden and less intuitively apparent. Although hillshading and image simulation methods described above can detect local

errors, the errors detected on the simulated images need to be traced back to the source data through visual inspection of source data overlaid on the top of the simulated images. This interactive process needs considerable effort, if the topographical data set under investigation is dense and huge. In this research, a new technique is developed for automatically detecting subtle local errors in irregularly distributed topographical data sets and incorporated in the IGIAS software package (Liu et al., 1998).

The fundamental assumption of this technique is the spatial continuity and autocorrelation of the terrain surface. The presence of serious errors tends to destroy the local continuity. Thus, checking the consistency and continuity for each data point in the context of nearby points can provide an important clue for error detection. If a point is strongly inconsistent with its neighbor points and statistically unreasonable, it is flagged as a local outlier in a statistical sense, meaning that the point has an anomalous and doubtful value.

To check local continuity and consistency, we need to define a neighborhood for each point by selecting a subset of surrounding points. To avoid the directional bias and keep the number of neighboring points from being excessively large or insufficiently small, we choose eight nearest points, each taken from one octant (Fig. 3.8). For each data point being checked, its octant neighbors are identified by using a super-block based sorting and searching scheme (Bentley et al., 1980; Al-Daoud and Roberts, 1996). The idea is to partition the irregularly distributed data set into an array of super blocks (cells or bins), then assign each point to a block according to its location. When processing a query point, we only need to search nearby blocks in a relatively small neighborhood of the query point instead of the entire data set (Fig. 3.9). A typical neighborhood configuration



⊕ Point in query o Unselected point

• Selected neighbor point

Fig. 3.8: Definition of octant neighbor points

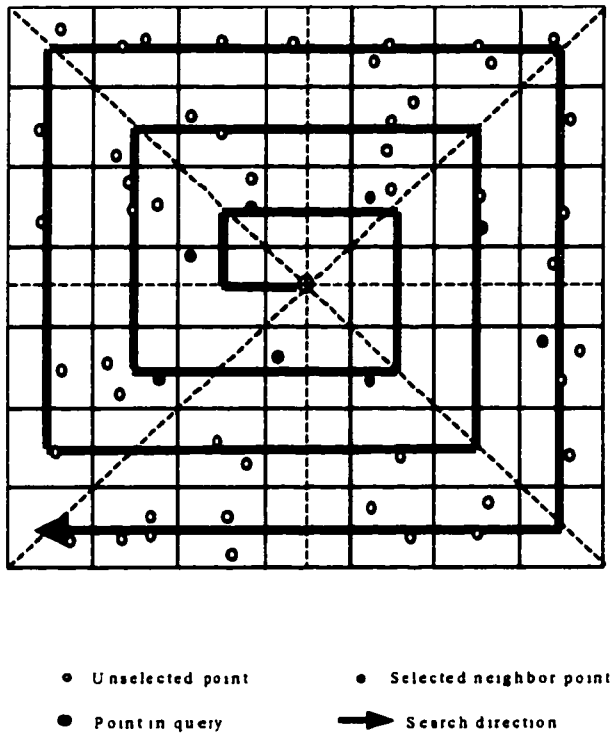


Fig. 3.9: Super-block based sorting and spiral searching

formed by octant neighbor points is illustrated in Fig. 3.10. Apparently, the size and shape of the octant neighborhood has the qualities of compactness and equiangularity. This makes the subsequent local consistency and continuity test more reliable and effective.

Based on octant neighbor points, we designed two local outlier indices: the residual index and the gradient index. With the continuity and autocorrelation assumption of the underlying surface, the elevation value of a data point can be well predicted through interpolation based on its surrounding neighboring points. The difference between the predicted value and the given value indicates the extent to which the data point under examination is an outlier with respect to its neighbors. To obtain a reliable prediction, we designed a robust IDW (Inverse Distance Weighted) interpolation algorithm, in which the predicted value is a linearly weighted function of its octant neighbors:

$$\hat{z}_q^* = \sum_{i=1}^{8} w_i z(p_i) \tag{3.10}$$

$$w_i = \frac{d_i^{-m}}{\sum_{j=1}^{8} d_j^{-m}} \tag{3.11}$$

Where $\hat{z}_q^*$ is the predicted elevation value for the query point q; $z(p_i)$ is the elevation value of neighbor point p_i; d_i is the distance between the query point and its neighbor point p_i; w_i is the weight of neighbor point p_i; and m is the distance friction factor.

One obvious problem with equation (3.10) and (3.11) is that if an outlier exists in the selected octant neighbor points, the prediction result will be contaminated and biased. To make the prediction robust in the face of outliers in neighbor points, we use a Jackknife

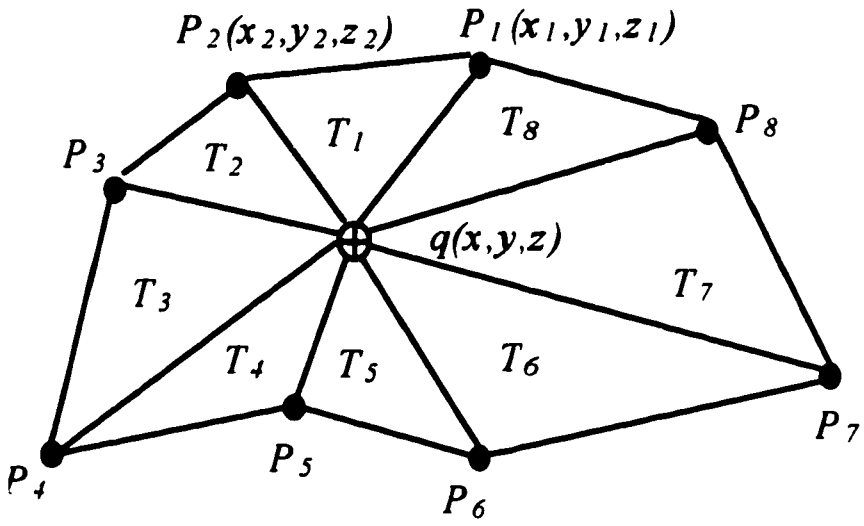


Fig. 3.10: A typical neighborhood configuration

technique (Burt and Barber, 1996). Namely, we drop one neighbor point at a time and use the remaining seven neighbor points to make the prediction. Repeating this procedure for every neighbor point, we obtain eight additional predictions:

$$\hat{z}_q^{(k)} = \sum_{i \neq k} w_i z(p_i) \quad (k=1, 2, \ldots, 8) \tag{3.12}$$

$$w_i = \frac{d_i^{-m}}{\sum_{j=1}^{8} d_j^{-m} - d_k^{-m}} \quad (k=1, 2, \ldots, 8) \tag{3.13}$$

Where $\hat{z}_q^{(k)}$ is the predicted elevation value with the neighbor point p_k omitted. Other parameters are the same as in equations (3.10) and (3.11). To obtain a representative and reliable surface elevation prediction at point q, we take the median value of these predicted values:

$$\hat{z}_q = median\{\hat{z}_q^*, \hat{z}_q^{(1)}, \hat{z}_q^{(2)}, \ldots, \hat{z}_q^{(8)}\} \tag{3.14}$$

The residual between the observed value (recorded in the data set) and the predicted value at point q is used as a local outlier index:

$$\Delta z_q = z_q - \hat{z}_q \tag{3.15}$$

Where Δz_q is the residual, and z_q is the observed elevation value at the query point q.

Surface gradient, indicating the surface shape and the local variation, is another good measure of the local surface continuity. The gradient is very sensitive to noise and errors. Unreasonable and erratic gradient values are often associated with serious errors in the elevation data. For an individual data point, the gradient can be estimated from its adjacent neighbor points. As shown in Fig. 3.10, the query point q in conjunction with its surrounding neighbor points form eight triangles. These triangles have various

orientations in 3D space, depending on the relative positions and heights of each pair of neighbors. The gradient for each triangle can be calculated by taking the cross product of any pair of sides from each triangle (Watson, 1992). For the triangle T_1 in Fig. 3.10, the cross product vector (X_1, Y_1, Z_1) is given by the Cartesian coordinates of the vertices $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$, and $q(x, x, z)$:

$$\begin{aligned} X_1 &= (y_2 - y_1)(z - z_1) - (y - y_1)(z_2 - z_1) \\ Y_1 &= (z_2 - z_1)(x - x_1) - (z - z_1)(x_2 - x_1) \\ Z_1 &= (x_2 - x_1)(y - y_1) - (x - x_1)(y_2 - y_1) \end{aligned} \tag{3.16}$$

This three-dimensional vector is perpendicular to the triangle plane T_1 and its length is twice the area of the triangle. The gradient G_1 and the area a_1 of the triangle plane T_1 can be calculated:

$$G_1 = \sqrt{\left(\frac{X_1}{Z_1}\right)^2 + \left(\frac{Y_1}{Z_1}\right)^2} \tag{3.17}$$

$$a_1 = \frac{\sqrt{X_1^2 + Y_1^2 + Z_1^2}}{2} \tag{3.18}$$

The weighted sum of the gradients of the surrounding triangles reflects the average gradient around the query data point q. For a large triangle, two neighbor points are far away from the query point, and the calculated gradient is less reliable for representing the surface shape around the query point. Therefore, the larger the triangle, the smaller the weight it receives. Thus the weighted average gradient can be defined as:

$$G_q^* = \sum_{i=1}^{8} G_i w_i \tag{3.19}$$

$$w_i = \frac{a_i^{-1}}{\sum_{j=1}^{8} a_j^{-1}} \tag{3.20}$$

where G_q^* is the weighted average gradient; G_i is the gradient of triangle T_i; a_i is the area of triangle T_i; w_i is the weight of triangle T_i.

The surface gradient calculated in this way is subject to error, if the outliers exist in the neighbor points. To make the gradient estimate robust to the presence of errors in the neighboring points, we again use the Jackknife technique. We subtract one triangle from the weighted sum at each time:

$$G_q^{(k)} = \sum_{i \neq k} w_i G_i \quad (k=1, 2, \ldots, 8) \tag{3.21}$$

$$w_i = \frac{a_i^{-1}}{\sum_{j=1}^{8} a_j^{-1} - a_k^{-1}} \quad (k=1, 2, \ldots, 8) \tag{3.22}$$

where $G_q^{(k)}$ is the weighted average gradient with the triangle T_k dropped out. We take the median of all these estimates as the final surface gradient around point q:

$$G_q = median\{G_q^*, G_q^{(1)}, G_q^{(2)}, \ldots, G_q^{(8)}\} \tag{3.23}$$

The residual index Δz_q and gradient index G_q make anomalous points more pronounced in the local context. To differentiate outliers from the rest of the normal points, we developed statistical criteria with reference to the local trend, namely, the central tendency and dispersion of the residuals and surface gradients in a small local area. Super blocks are used as the basic unit to construct a local area for estimating the local central tendency and dispersion. The minimum number of data points in the local area is set to be 30. For each non-empty super block, we check the number of data points

inside. If the number is smaller than 30, the local area is expanded in a spiral-like fashion (Fig. 3.9) by including the surrounding super blocks until the number of points equals or is greater than 30.

The residual Δz_q has signed values, indicating the difference between the given value and the predicted value from nearby points. If the point being checked is contaminated by serious error, the absolute value of Δz_q will be significantly larger and will markedly deviate from the central tendency. If the calculated residual is beyond a specified confidence interval of the central tendency, the point under investigation is declared as an outlier. The sample median is used as the robust estimator for the central tendency and the median deviation (Hampel, 1974; Barnett and Lewis, 1984) as the robust estimator for the spread of the local sample:

$$\begin{aligned} \Delta\tilde{z} &= median\{\Delta z_j\} \qquad (j = 1,2,...,m) \\ s_{\Delta z} &= median\{\left|\Delta z_j - \Delta\tilde{z}\right|\} \qquad (j = 1,2,...,m) \end{aligned} \tag{3.24}$$

The distribution of the deviation/spread statistic to be used in the confidence limit calculation is the Student t distribution with m-1 degrees of freedom (Barnett and Lewis 1984; Burt and Barber, 1996), and a local confidence interval can be constructed at any desired level of confidence 1-α:

$$(\Delta\tilde{z} - t_{\alpha/2,m-1} * \frac{s_{\Delta z}}{\sqrt{m}},\ \Delta\tilde{z} + t_{\alpha/2,m-1} * \frac{s_{\Delta z}}{\sqrt{m}}) \tag{3.25}$$

$t_{\alpha/2,m-1}$ refers to the value of statistic t with an upper-tail probability of $\alpha/2$ and with m-1 degrees of the freedom. The median and median deviation provide strong protection against the influence of extreme outliers for estimating the central tendency and dispersion of the underlying distribution. Through the use of these two robust estimators,

we obtain a robust confidence interval and t-test over a spectrum of possible distributional forms, such as a contaminated version of normal or other symmetric distributions. The statistical discordancy test can be performed by examining whether Δz_q, the residual value at point q, is within the confidence interval. Equivalently, we can compare t value at point q calculated in (3.26) with known significance levels $t_{\alpha/2,m-1}$:

$$t_q = \frac{\left|\Delta z_q - \Delta\tilde{z}\right|}{s_{\Delta z}/\sqrt{m}} \tag{3.26}$$

where t_q is the t value calculated for the query point q; and m (>=30) is the number of data points in the defined local area. If $t_q > t_{\alpha/2,m-1}$, the residual value at point q would be out of the confidence interval calculated in (3.25), and we can declare that the point q is an outlier because it is statistically unreasonable and suspicious at the level of test.

For gradient G_q, its distributional form in the defined local area might be more likely away from the normal distribution. Hence, the use of robust estimators is more important for obtaining reliable estimates for the central tendency and dispersion. Similarly, the median and the median deviation of the gradients in a local area are used to represent the central tendency and dispersion:

$$\begin{aligned} \tilde{G} &= median\{G_j\} \qquad (j = 1,2,...,m) \\ s_G &= median\{\left|G_j - \tilde{G}\right|\} \qquad (j = 1,2,...,m) \end{aligned} \tag{3.27}$$

Since the gradient G_q has positive values and we are only interested in the upper bound outliers, the one-sided t test is conducted:

$$t_q = \frac{\left|G_q - \tilde{G}\right|}{s_G/\sqrt{m}} \tag{3.28}$$

If $t_q > t_{\alpha,m-1}$, namely, the gradient value at point q is out of the confidence interval, then we declare the point q as an outlier.

The combination of a super-block based searching strategy and locally adaptive and robust statistics makes this technique both computationally efficient and functionally effective in detecting local anomalous values in topographical data sets. The implementation details can be found in Liu et al. (1999). An early version of this algorithm, which has similar attributes except for the use of global central tendency and dispersion as statistical criteria for identifying outliers, has been applied to ERS satellite radar altimeter data of the Antarctic. Numerous local errors have been detected and removed. The full version of this algorithm was successfully applied to airborne radar data over the Evans Ice Stream and an irregularly distributed elevation data set derived from SAR stereo technique over the Terra Nova Bay area. Fig. 3.11 illustrates the erroneous data points detected in the Evans data set. The error detection results for the SAR stereo derived elevation data set will be presented in Chapter 7.

3.6 Summary

At the beginning of this chapter, We analyzed the comparative advantages and disadvantages of different topographical data acquisition techniques and highlighted the need, potential benefits and problems for integrating various types of topographical measurements. We pointed out that the construction of a consistent and error-free topographical database is the essential prerequisite for the integration of diverse topographical source data. In the case of the Antarctic, we demonstrated that the available data import and conversion functions in the proprietary GIS software can ingest

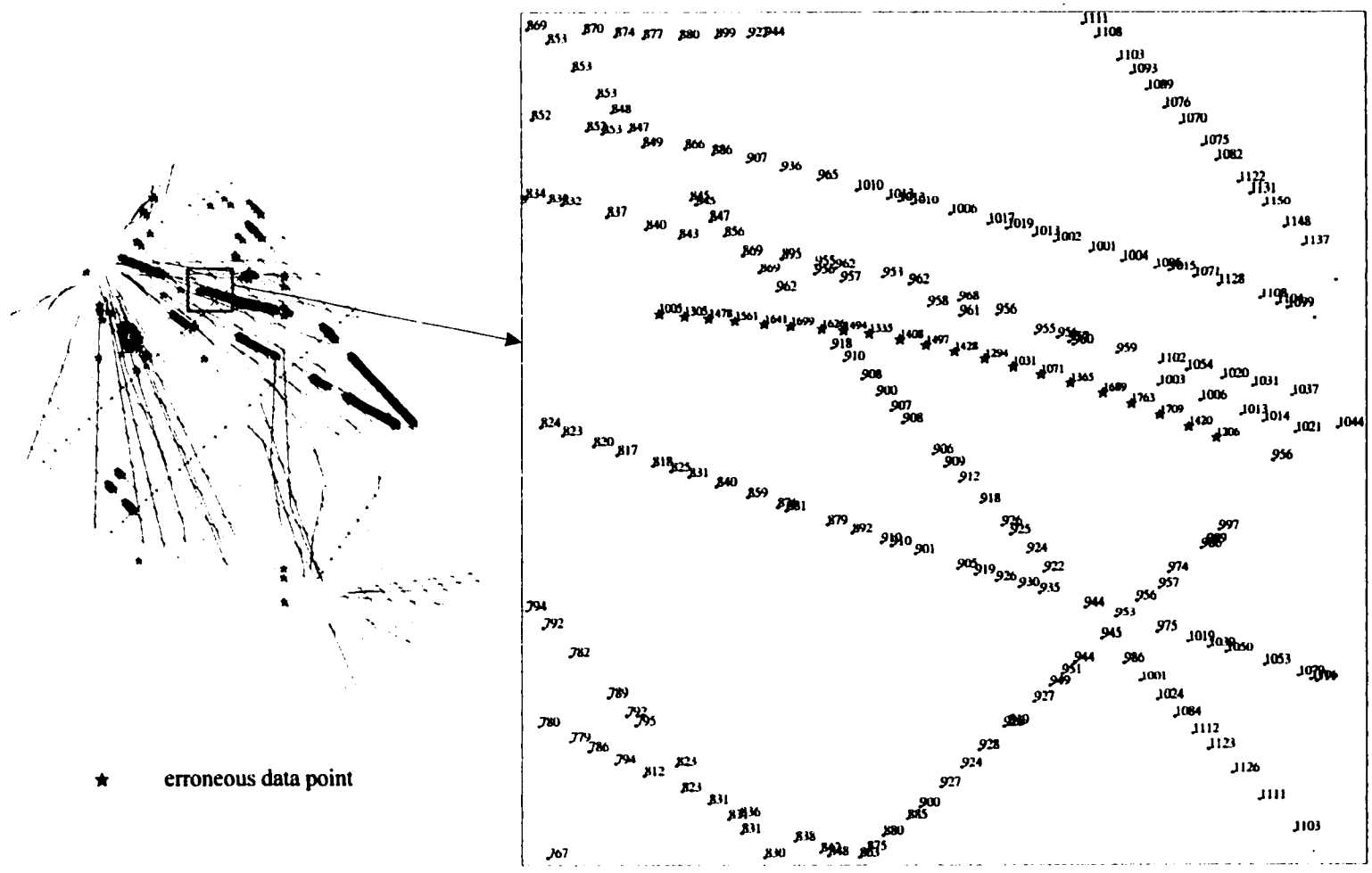


Fig.3.11: Erroneous data points in Evans RES data set detected using locally adaptive and robust statistical analysis method

various types of topographical data into coherent, compatible, and topologically structured formats-vector coverages or grid layers, which can be easily and efficiently stored, retrieved and displayed in an unified GIS environment. The mathematical operations on attribute tables and comprehensive projection functions provided in the GIS software enable us to geocode, co-register and interrelate different topographical data layers using a common planimetric and vertical reference systems. Nevertheless, the available GIS software does not have any explicit error checking and correction functions. To fill this gap, we developed a comprehensive error detection and correction scheme. Based on the ARC/INFO internal functions, we developed global statistical analysis, visualization and cross-validation methods. Using C language programming, we implemented a technique for producing a synthetic stereo image for visualizing and checking the gross errors in the source data. We also implemented a rigorous SAR image simulation technique to detect and correct positional errors in source data that proved very effective. A new local algorithm is developed for detecting subtle local errors in irregularly distributed spatial data sets based on super-block based searching and locally adaptive and robust statistical criteria. Its effectiveness and sharpness has been demonstrated on the Evans data set and will be further shown later in Chapter 7.

Since the topographical source data of the Antarctic have undergone rigorous and extensive error detection and correction procedures as described above, We believe with a high level of confidence that the overwhelming majority of outliers and errors in the source data have been corrected or removed. The resulting Antarctic topographical database is reliable and free of gross errors and blunders.

CHAPTER 4

SPATIAL INTERPOLATION OF TOPOGRAPHICAL DATA

4.1 Overview of spatial interpolation algorithms

Spatial interpolation is a computational procedure that estimates surface values at unsampled positions within the area covered by a sample of existing measurements. The primary computation involved in the large-scale DEM generation is to interpolate the topographical source data, either irregularly distributed point measurements or contour lines, into a grid of elevation values with a specified origin, orientation and grid spacing (cell size). This computational procedure is also known as gridding. In some cases, an elevation grid needs to be interpolated into a new regular elevation grid with a different origin, orientation and/or grid spacing. This grid-to-grid interpolation process is referred to as resampling.

Numerous spatial interpolation algorithms have been developed in the past. According to the spatial extent of the input data points involved in the calculation, the spatial interpolation algorithms can be divided into global methods and local methods (Lam, 1983; Burrough 1986). Global interpolation methods exploit all the sample points in estimating elevation values at new positions. They are often used to model the regional trends (long range variation). Examples include bivariate low-order (linear,

quadratic, or cubic) polynomial regression and Fourier series. With global interpolation methods, local variations are often suppressed and the computation involved is very intensive. The requirements on the computation efficiency and the representation of local topographical details preclude the global interpolation methods for this research.

In recognition of the fact that the elevation values at distant locations have little similarity and correlation, local interpolation methods only utilize nearby points within a limited range of the surrounding region. Local interpolation algorithms can be further classified into two approaches: fitted functions and weighted averages (Watson, 1992). The fitted function algorithms partition data points into small patches and then determine the coefficients of a selected analytical bivariate function based on the sample points inside each patch, frequently using least squared criteria. With the fitted coefficients, the elevation value at a given position can be obtained through the evaluation of the function. Fitting a piecewise function on the patch has a smoothing effect because some degree of local details may be submerged in the overall local trend surface. The order of the analytical function, and hence the number of coefficients, determines the shape and complexity of the local surface. On the other hand, weighted average interpolation algorithms determine the elevation value at each interpolation point by a weighted summation of the elevation values of a running subset of neighboring data points, without using an intermediate parametric surface.

Based on internal functions of the proprietary GIS software ARC/INFO, the TIN linear, TIN quintic, ordinary Kriging, universal Kriging, minimum curvature spline and the TOPOGRID interpolation algorithms were incorporated in the IGIAS. Using the C-programming language, we implemented directional Inverse Distance Weighted (IDW)

and linear gradient descent interpolation methods. All these algorithms belong to the category of local interpolation methods. Among them, IDW and Kriging use the weighted average approach, while TIN linear, TIN quintic, and minimum curvature spline employ the fitted function approach. The TOPOGRID algorithm follows an iterative finite difference interpolation procedure, which is derived using a variational approach under the surface smoothness constraint. The linear gradient descent algorithm is designed for interpolating rasterized contour data. All these alternative interpolation algorithms were integrated under the Graphical User's Interface (GUI) of the IGIAS. Through the GUI, we can easily access each of the algorithms and perform interpolation operations in a GIS environment. Input source data, either line-type or point type data, can be conveniently preprocessed and loaded from the topographical database, and interpolation parameters and areal extent can be set in the graphical menu (Fig. 2.2). The interpolation result is then visualized and compared with the original source data and actual satellite images in various ways, as will be demonstrated later in this chapter and Chapter 6. The evaluation results from visualization and comparison are used as important feedback information for adjusting interpolation parameters or selecting an appropriate interpolation algorithm.

Despite a variety of interpolation algorithms, this research shows that there is no best single algorithm that is absolutely superior to all others and equally suitable for all situations, as concluded by others (Lam, 1983; Weibel and Heller, 1991). It is found that the performance of different algorithms is strongly dependent on the pattern, density and format of source data, as well as the complexity of the terrain surface under investigation. In general, when topographical measurements are abundant and evenly distributed, most

interpolation techniques give similar and satisfactory results (Weber and Englund, 1992; Declercq, 1996). When the input data are sparse and unevenly distributed, however, the choice of algorithm and its parameters can be critical for obtaining a plausible interpolation result. In this research, the evaluation and selection of interpolation techniques are based on three considerations:

- The elevation values and pattern of the interpolated surface should be in good agreement with the source data. At the locations coincident with or in the vicinity of the source data points or contour lines, the resulting elevation grid should have identical or close values as the source data. In the area bounded by a contour pair, the interpolated values must be within the range defined by the two contour labels, and should gradually increase from the down-slope contour line to the up-slope contour line. The terrain trend represented by the interpolated surface should be consistent with that expressed by the source data. For example, the location and orientation of hills, basins, ridges and valleys evident in contour data should be faithfully reflected by the interpolated surface. This can be verified by comparing the contour lines derived from the interpolated elevation grid with the original input contour lines. The consistency checking ensures that the selected interpolation method does not distort the source data, and the topographical details present in the source data can be transferred into the resulting elevation grid.
- The interpolated surface should be single-valued, continuous, and smooth at all positions where no evidence supports the existence of surface breaklines. Practically, this property is checked by the visual inspection of the plausibility of the resulting surface. The hillshaded image or simulated stereo image of the interpolated elevation

grid must be artifact-free, and the shape of terrain features in the simulated image should agree with that suggested by actual images.

- The computation involved must be efficient and fast, and the computer memory requirement should be limited in view of the large amount of data being processed.

The assembled topographic data have two distinct formats: line-type contour data, and point-type survey and radar altimeter data. The pattern and density of point-type data vary considerably. The preprocessed satellite altimeter data are evenly distributed in a quasi-grid format, while ground survey data and GPS data are sparsely scattered over space. Airborne radar data are anisotropically distributed, namely, densely sampled along flight lines but widely separated between flight transects. We tested different interpolation algorithms on each type of data and corresponding terrain, and identify the drawbacks of the existing algorithms. Some new and improved interpolation procedures are designed for contour data and traverse data.

4.2 Interpolation of evenly distributed point data

When input point measurements are evenly distributed over space, the surface is well controlled, and there is no need to account for the clusters and anisotropic pattern of source data. The ERS-1 satellite radar altimeter data that we used in the final DEM generation were preprocessed in the spherical geographical coordinate system (latitude and longitude). When this preprocessed data set is projected into the polar stereographic projection, the data points are evenly distributed, but not in regular grid (Fig. 4.1). The spacing between data points is approximately 5 km. It should be noted that before the

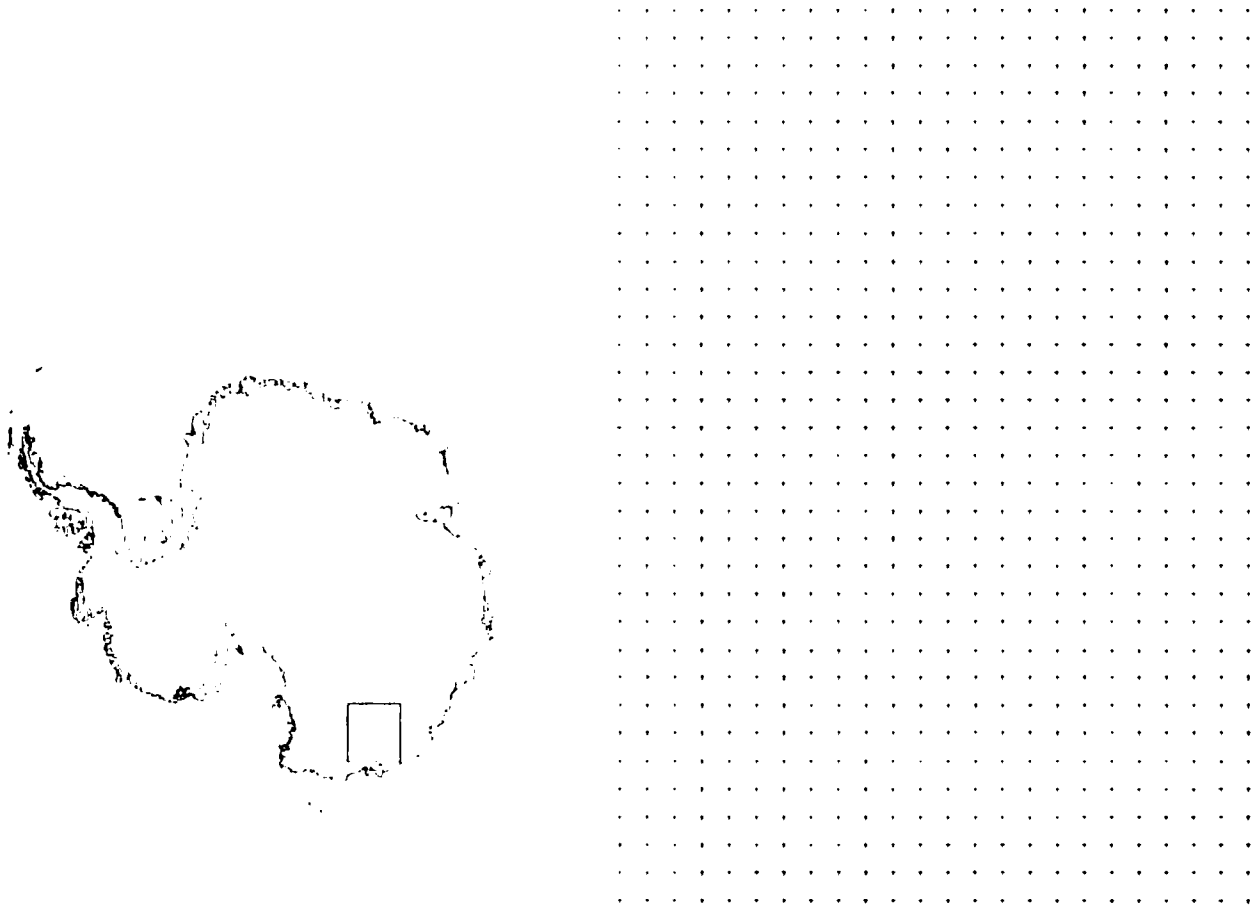


Fig. 4.1: Spatial distribution pattern of preprocessed ERS-1 Radar altimeter data in the Victoria Land, Antarctica

preprocessing the original satellite radar altimeter measurements are distributed along satellite tracks rather than evenly distributed.

Kriging, minimum curvature spline and TIN-based methods were tested on a test site in Victoria Land, Antarctica. Kriging is an elegant and optimal geo-statistical interpolation algorithm (Cressie, 1993). Ordinary Kriging assumes that the surface variation is statistically stationary and free of any systematic trend, while universal Kriging accounts for a structural component (drift), namely, a constant trend over the surface. Kriging uses a weighted linear combination of a number of neighboring sample values to model the spatial variation within a local area bounded by the input sample points (Bailey and Gatrell, 1995; Isaaks and Srivastava, 1989):

$$\hat{Z}(s) = \sum_{i=1}^{n} w_i(s) Z(s_i) \tag{4.1}$$

$$\sum_{i=1}^{m} w_i = 1 \tag{4.2}$$

Where $\hat{Z}(s)$ is the estimated value at location s, $Z(s_i)$ is the observed value at location s_i; n is the number of points used in the calculation; and $w_i(s)$ is the weight of the data point at s_i.

The weight $w_i(s)$ is optimally determined such that the estimated elevation value is unbiased, and that the estimation variance is less than for any other linear combination of the observed values. Mathematically, the weights are computed by using the covariance matrix (Isaaks and Srivastava, 1989):

$$w = C^{-1} D \tag{4.3}$$

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} \quad C = \begin{bmatrix} \tilde{C}_{11} & \cdots & \tilde{C}_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \tilde{C}_{n1} & \cdots & \tilde{C}_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \quad D = \begin{bmatrix} \tilde{C}_{10} \\ \vdots \\ \tilde{C}_{n0} \\ 1 \end{bmatrix} \tag{4.4}$$

where μ is a constant to be determined; the matrix C consists of covariance values between sample points, accounting for the possible information redundancies between sample points, namely, the adverse effect of the clustering of samples, and the vector D consists of covariance values between the sample points and the point being estimated, accounting for the spatial proximity of the sample points to the estimate point in terms of statistical distance. The minimized error variance, i.e., Kriging variance can be calculated (Isaaks and Srivastava, 1989) as:

$$\sigma^2 = \sigma_0^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \tilde{C}_{ij} - 2\sum_{i=1}^{n} w_i \tilde{C}_{i0} \tag{4.5}$$

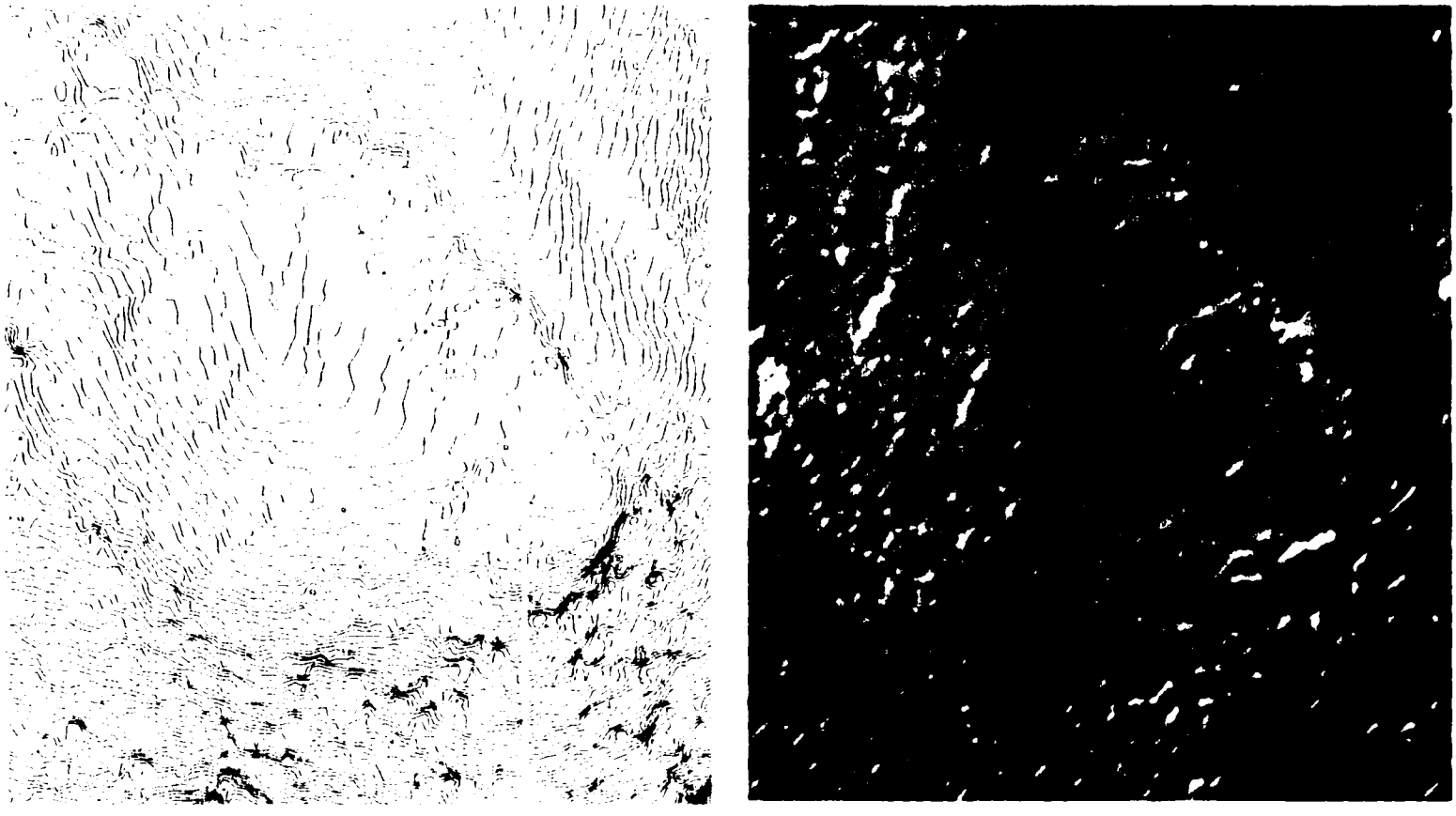
where σ_0^2 is the nugget effect due to the original measuring error and small-scale variation; $\tilde{C}_{ij}$ is the covariance between sample point i and sample point j; and $\tilde{C}_{i0}$ is the covariance between sample point i and the point being estimated. Obviously, the magnitude of estimation variance largely reflects the degree of spatial autocorrelation and geometric configuration (density and pattern) of sample points with respect to the point being estimated. The covariance matrix C and D, describing the basic structure of spatial dependence of elevation values, is modeled by continuous theoretical functions such as spherical, exponential, and Gaussian functions. The parameters of the covariance function can be specified using a prior knowledge or fitted based on an empirical variogram constructed by sample points. An empirical variogram is defined as:

$$\hat{\gamma}(h) = \frac{1}{2m} \sum_{i=1}^{m} \{Z(x_i) - Z(x_i + h)\}^2 \qquad (4.6)$$

Where $\hat{\gamma}(h)$ is the variogram value at lag distance h; and m is the number of pairs of sample points separated by a distance h.

Since the regional trend is obvious in the terrain surface, universal Kriging was used, instead of ordinary Kriging. As the inverting of the covariance matrix is involved, the Kriging method is computationally very intensive, proportional to the cube of the number of data points. The interpolation result with universal Kriging on the test site of Victoria Land is presented in Fig. 4.2.

Among many variants of spline interpolation, two-dimensional minimum curvature spline is analogous to a thin elastic plate passing through each of the data points with a minimum amount of bending. Minimum curvature spline generates the smoothest possible surface while attempting to honor the sample points as closely as possible. The interpolation function is derived by minimizing the surface curvature represented by the second order derivative terms using a variational approach. Rapid changes in gradient (the first derivative) may occur in the vicinity of data points, and the main drawback of the basic minimum curvature spline interpolation is the generation of overshoots in regions with a rapid change of gradient. By including first derivative terms in the minimization criteria, the stiffness of the interpolation plate can be reduced, and hence the overshoot problem can be alleviated. This gives rise to the thin plate spline with tension. The third derivative terms can be also incorporated into the minimization criteria to obtain a smoother surface. The general minimum curvature spline interpolation (Franke, 1982, 1985; Mitas and Mitasova, 1988; Mitasova and Mitas, 1993; Mitasova and



(a) (b)

Fig. 4.2 Interpolation result using universal Kriging on the evenly distributed data set.
(a)contours derived from the interpolated DEM; (b) Hill-shaded image of the interpolated DEM.

Hofierka, 1993; ESRI, 1991) uses the following functions in the ARC/INFO implementation:

$$Z(x,y) = T(x,y) + \sum_{j=1}^{N} \lambda_j R(r_j) \tag{4.7}$$

where N is the number of sample points used; λ_j are coefficients found by the solution of a system of linear equations; and r_j is the distance from the point being estimated to the jth sample point. T(x,y) and $R(r_j)$ are defined differently depending upon the selected option. The REGULARIZED option incorporates third derivative terms into the minimization criteria and obtains:

$$T(x,y) = a_1 + a_2 x + a_3 y \tag{4.8}$$

$$R(r) = \frac{1}{2\pi}\{\frac{r^2}{4}[\ln(\frac{r}{2\tau}) + c - 1] + \tau^2[K_0(\frac{r}{\tau}) + c + \ln(\frac{r}{2\pi})]\} \tag{4.9}$$

and the TENSION option incorporates first derivative terms into the minimization criteria and obtains:

$$T(x,y) = a_1 \tag{4.10}$$

$$R(r) = \frac{1}{2\pi\varphi}[\ln(\frac{r\varphi}{2}) + c + K_0(r\varphi)] \tag{4.11}$$

where τ^2 and φ^2 are the parameters controlling the weight of the third derivative term and first derivative term in the minimization criteria; r is the distance between the point being estimated and the sample point; K_0 is the modified Bessel function; c is a constant equal to 0.577215; and a_i are coefficients found by the solution of a system of linear equations. For computational purposes, the entire space of the output grid is divided into blocks or regions equal in size.

Both REGULARIZED and TENSION options were tested on the Victoria Land data set, and the results are similar for this test site. Fig. 4.3 shows the interpolation result of the minimum curvature spline with regularized option.

TIN-based interpolation is a two-step procedure: triangulation of sample points and locally fitting polynomials to triangles. Among many possible triangulation schemes, the Delaunay triangulation is commonly used. The resulting triangles are nearly equi-angular, reducing potential problems of long skinny triangles. The pure Delaunay criterion is often relaxed to accommodate surface breaklines. The triangulation process is constrained so that edges of triangles go along surface discontinuities instead of crossing linear discontinuity. Once the TIN has been constructed, linear planar or Quintic surfaces can be fitted to relevant triangle patches, and then elevation values at nodes of a grid can be estimated by evaluating the fitted linear or quintic functions. The generalized equation for linear interpolation of a point (x, y, z) in a triangle facet is:

$$Z=a+bx+cy \tag{4.12}$$

where *a, b, c* are fitted coefficients from three nodes of the triangle. Although the linear planar surface fitting on the triangles produces a continuous faceted surface, the surface slope changes abruptly when crossing over edges between adjacent triangles. The quintic interpolation uses a bivariate fifth-degree polynomial in x and y (Akima, 1978; ESRI, 1991):

$$Z(x,y)=\sum_{j=0}^{5}\sum_{k=0}^{5-j} q_{jk}\,x^{j}\,y^{k} \tag{4.13))}$$

where Z(x,y) is the interpolated value at the location (x, y); q_{jk} are the fitted coefficients, which are determined based on the heights of the triangle nodes as well as their first and

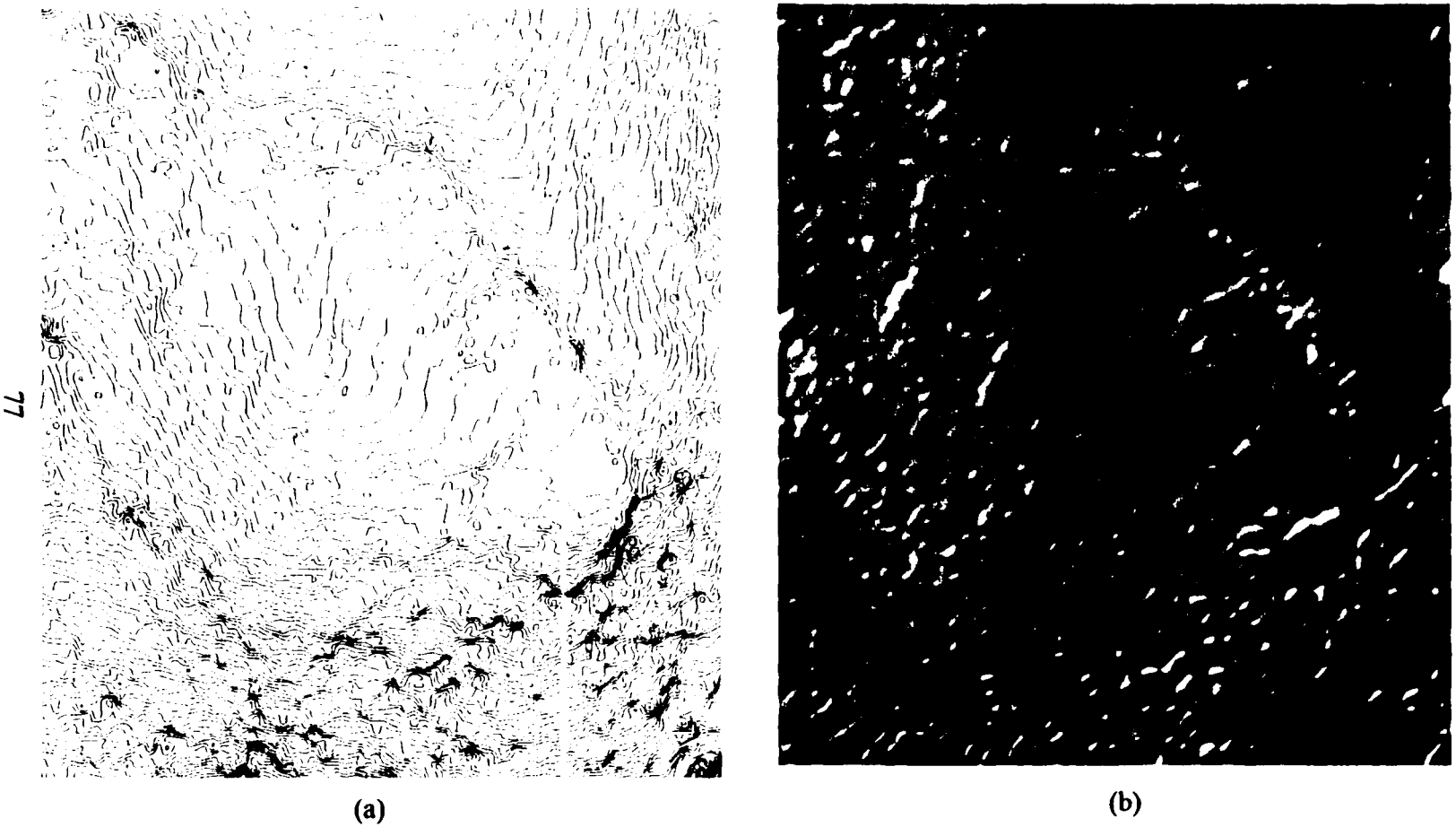


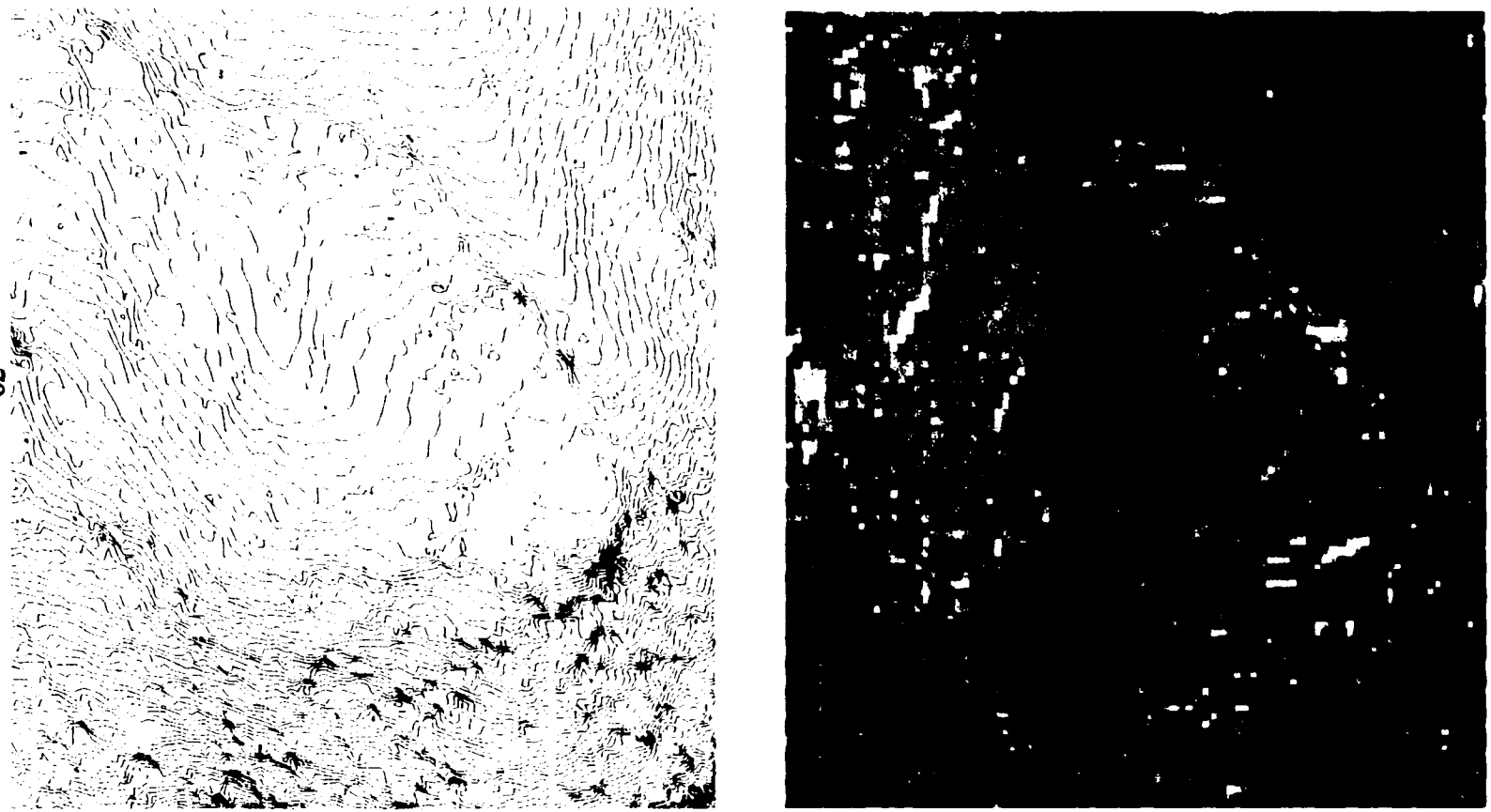
Fig. 4.3: Interpolation result using minimum curvature spline with regularized option on the evenly distributed data. (a) Contours derived from the interpolated DEM; (b) Hill-shaded image of the DEM.

second order derivatives. The derivatives at a node depend on the heights of all its neighboring points (Akima, 1978). The interpolation results with TIN linear and TIN quintic methods on the Victoria Land data set are shown in Fig. 4.4 and Fig. 4.5 respectively.

Inspection of the interpolation results shows that all these techniques give satisfactory results on this evenly distributed data set in terms of honoring the original measurements and depicting the terrain features. The mean of the differences between these interpolation results is zero, and the standard deviation of the differences is less than 3 m. Despite the similar interpolation results, the computational costs are quite different. The TIN-based method is computationally much more efficient than the Kriging method. It is also more efficient than the minimum curvature spline method. Therefore, the TIN algorithm is selected to interpolate the preprocessed satellite radar altimeter data. Between linear and quintic options, the TIN quintic interpolation algorithm is chosen. This is because the second derivatives of the surface from the TIN quintic interpolation are continuous and differentiable and the resulting surface looks more realistic (Fig. 4.5). Due to the gradient change between two adjacent triangles, the TIN linear interpolation algorithm gives a continuous but not a smooth surface (Fig. 4.4).

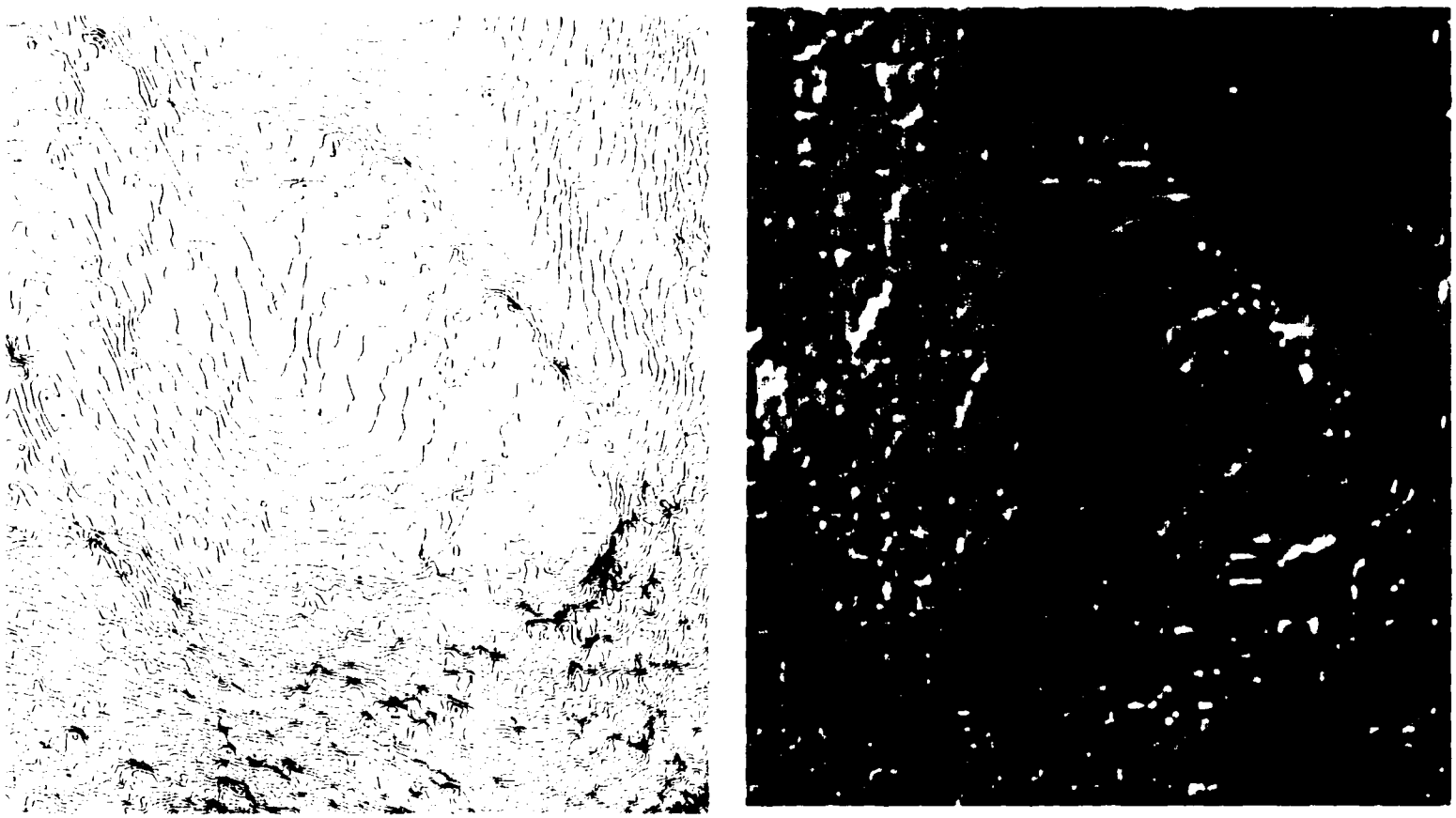
4.3 Interpolation of traverse data

With the development of laser and radar altimetry technology, an increasing volume of traverse topographical data have been collected by the airborne and spaceborne altimeters. Along flight tracks, the terrain surface is very densely sampled, and consequently some degree of data redundancy exits. On the other hand, the flight tracks



(a) (b)

Fig. 4.4: Interpolation result with TIN linear option on the evenly distributed data set.
(a) Contours derived from the interpolated DEM; (b) Hill-shaded image of the interpolated DEM.



(a) (b)

Fig. 4.5: Interpolation result with TIN quintic option on the evenly distributed data set.
(a) Contours derived from the interpolated DEM; (b) Hill-shaded image of the interpolated DEM.

are often widely separated, giving rise to large data gaps bounded by the flight tracks. For example, the airborne radar echo sounding data over the Evans Ice Stream and Fowler Peninsula have a sampling spacing of 50-90 m, but the distance between tracks is as large as 10 km (Fig. 4.6). Similarly, the airborne radar data over Ice Streams A, B, and C have a sampling spacing of about 120 m along track, and the gaps between tracks ranges from 5 to 10 km. Due to the anisotropic distribution of data density, traverse data impose serious difficulties on general-purpose interpolation algorithms.

As stated in the preceding section, all local interpolation algorithms need to define which sample points should be included in the calculation of a surface value at a specified grid node. Most general-purpose algorithms assume that the sample points are randomly or evenly distributed, and a simple rule is used to select an appropriate number of nearby points in the interpolation. If the selected points are too many, the computation will be very intensive, and the local details may be smoothed out. If the selected points are too few, the interpolation result will not be robust and stable. In the ARC/INFO implementation, Kriging and minimum curvature spline algorithms use the maximum or minimum number of nearest sample points and/or a specified search radius to define the selection criteria. The appropriate specification of the number of sample points and search distance can only avoid the problems associated with an excessively large or insufficiently small interpolation neighborhood. However, in the case of traverse data, either Kriging or the minimum curvature spline algorithm gives an unsatisfactory result (see Fig. 4.7 and Fig. 4.8). This is because the selected sample points tend to be from a small segment of the flight track. Due to the strong correlation and information redundancy between the selected sample points, the net information content that can be

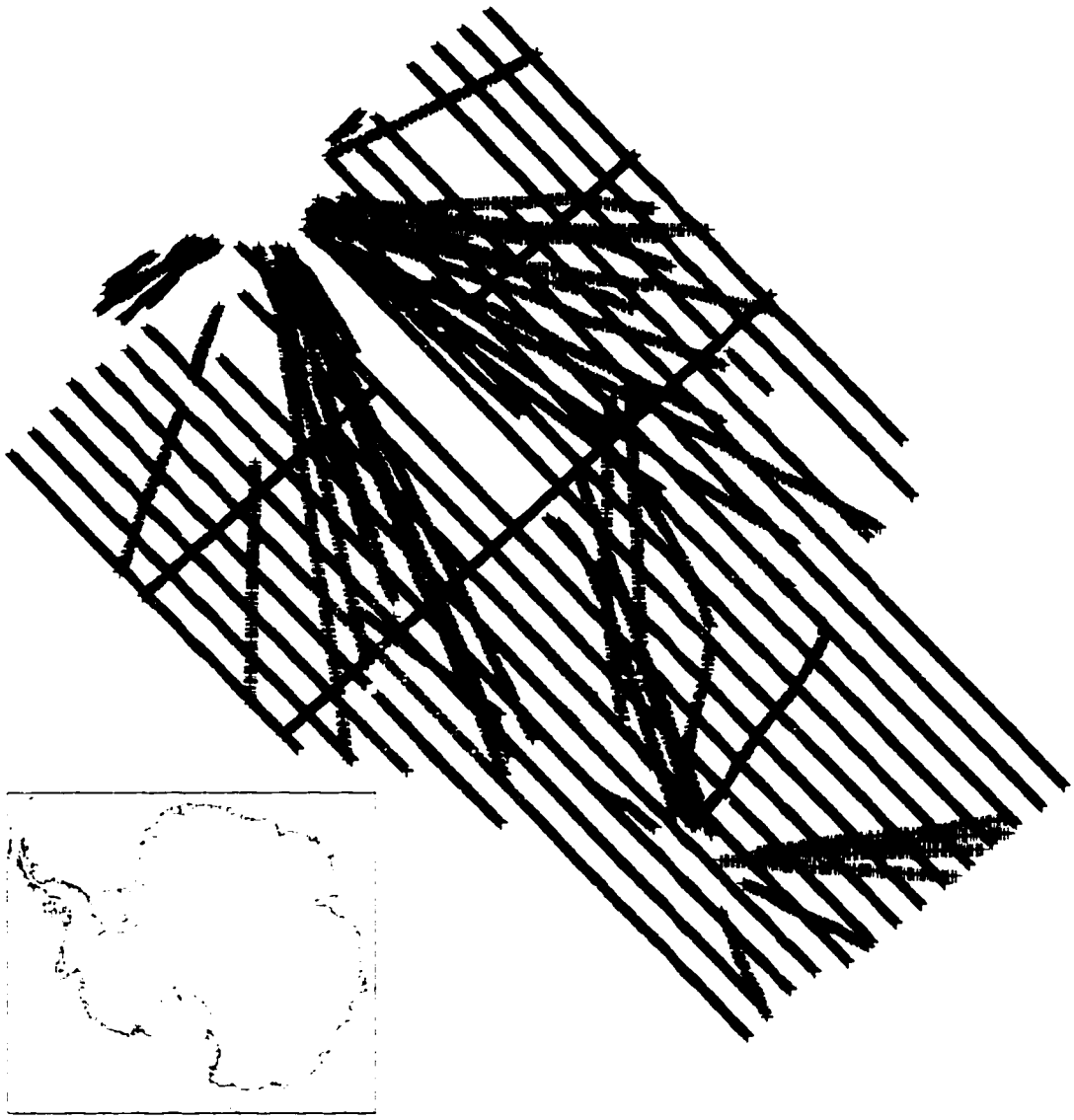
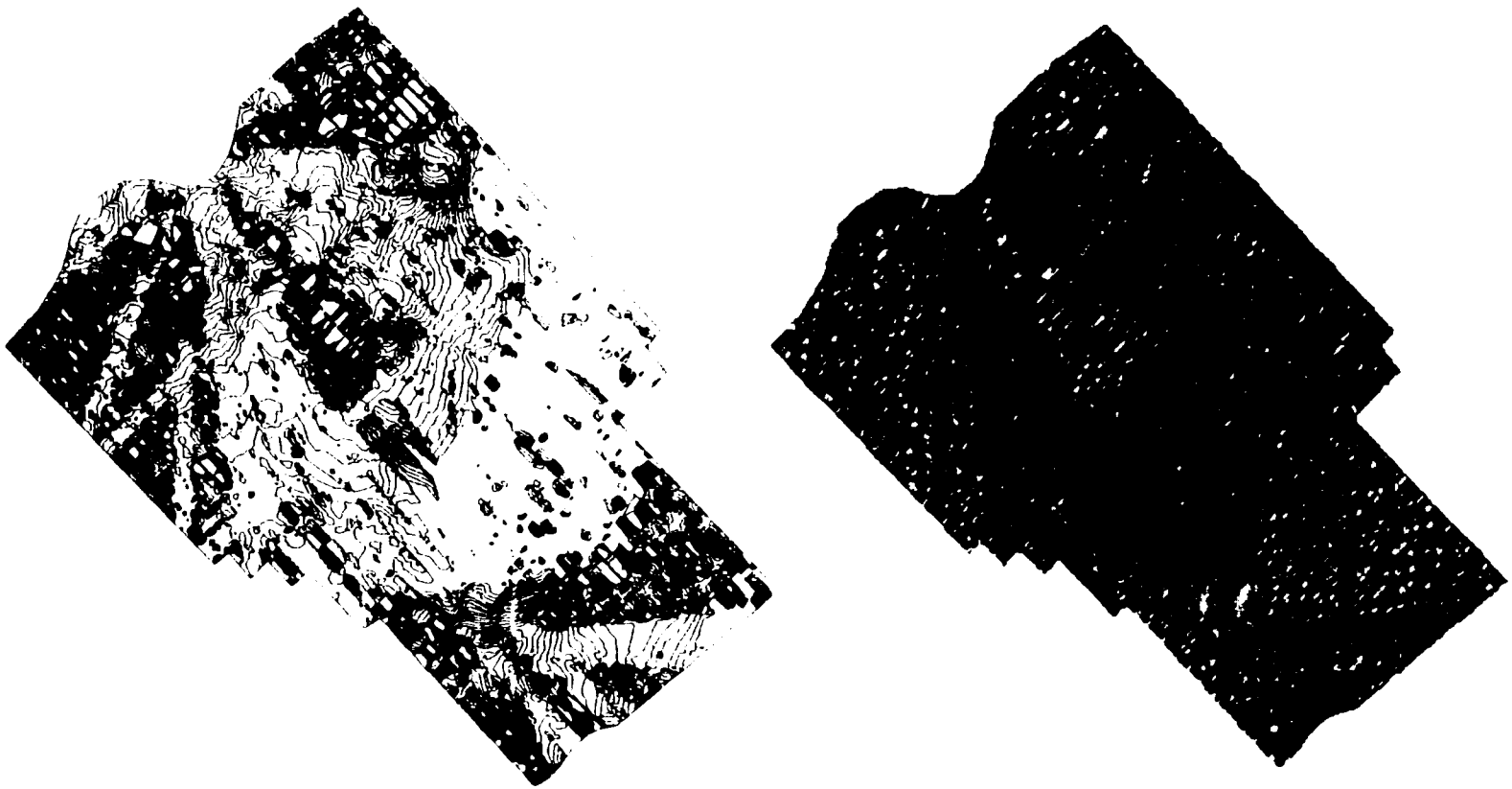
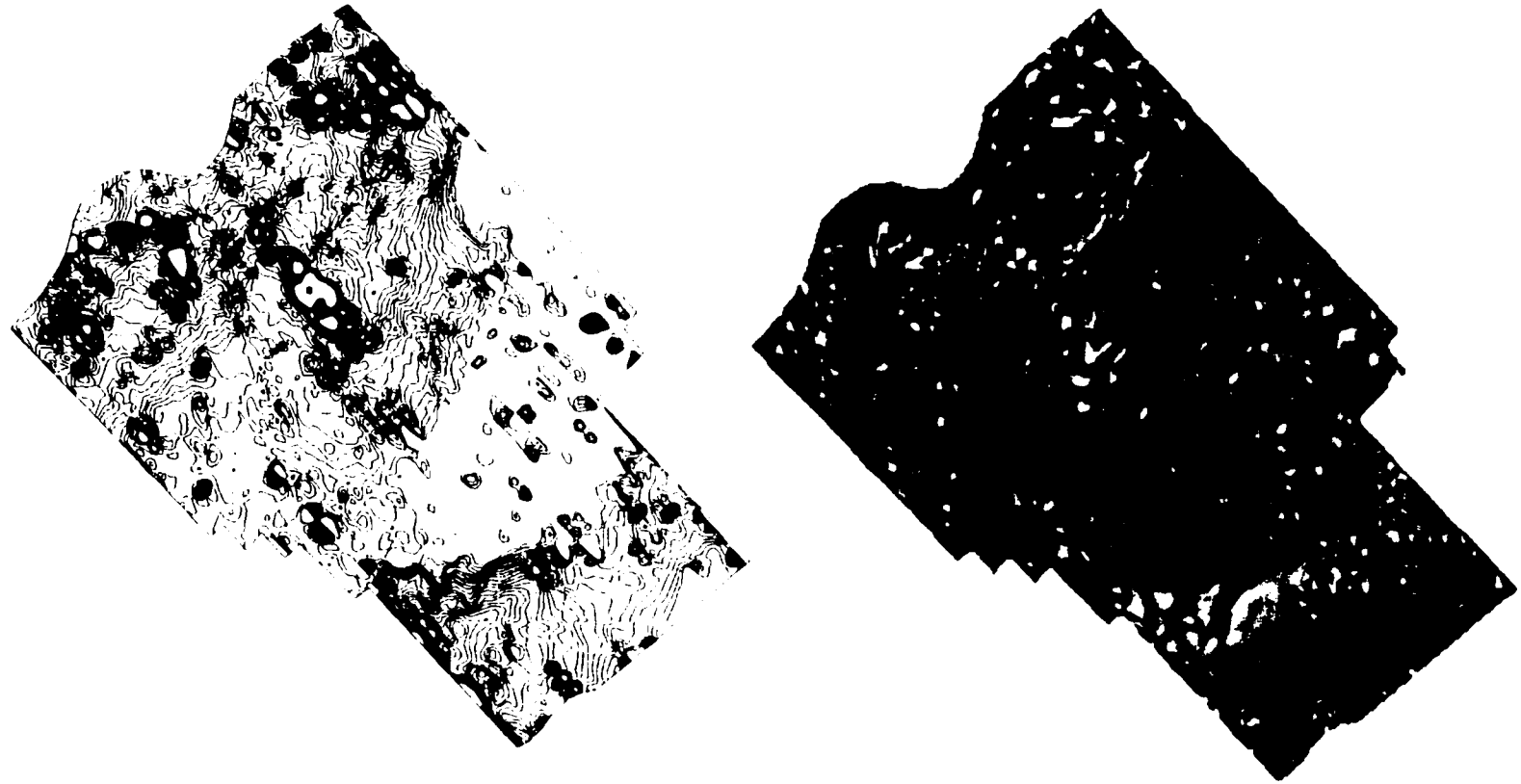


Fig. 4.6: Spatial distribution pattern of airborne traverse data over the Evans Ice Steam, Antarctica



(a) (b)

Fig. 4.7: Interpolation result with universal Kriging on the traverse data set.
(a) Contours derived from the interpolated DEM; (b) Hill-shaded image of the interpolated DEM.



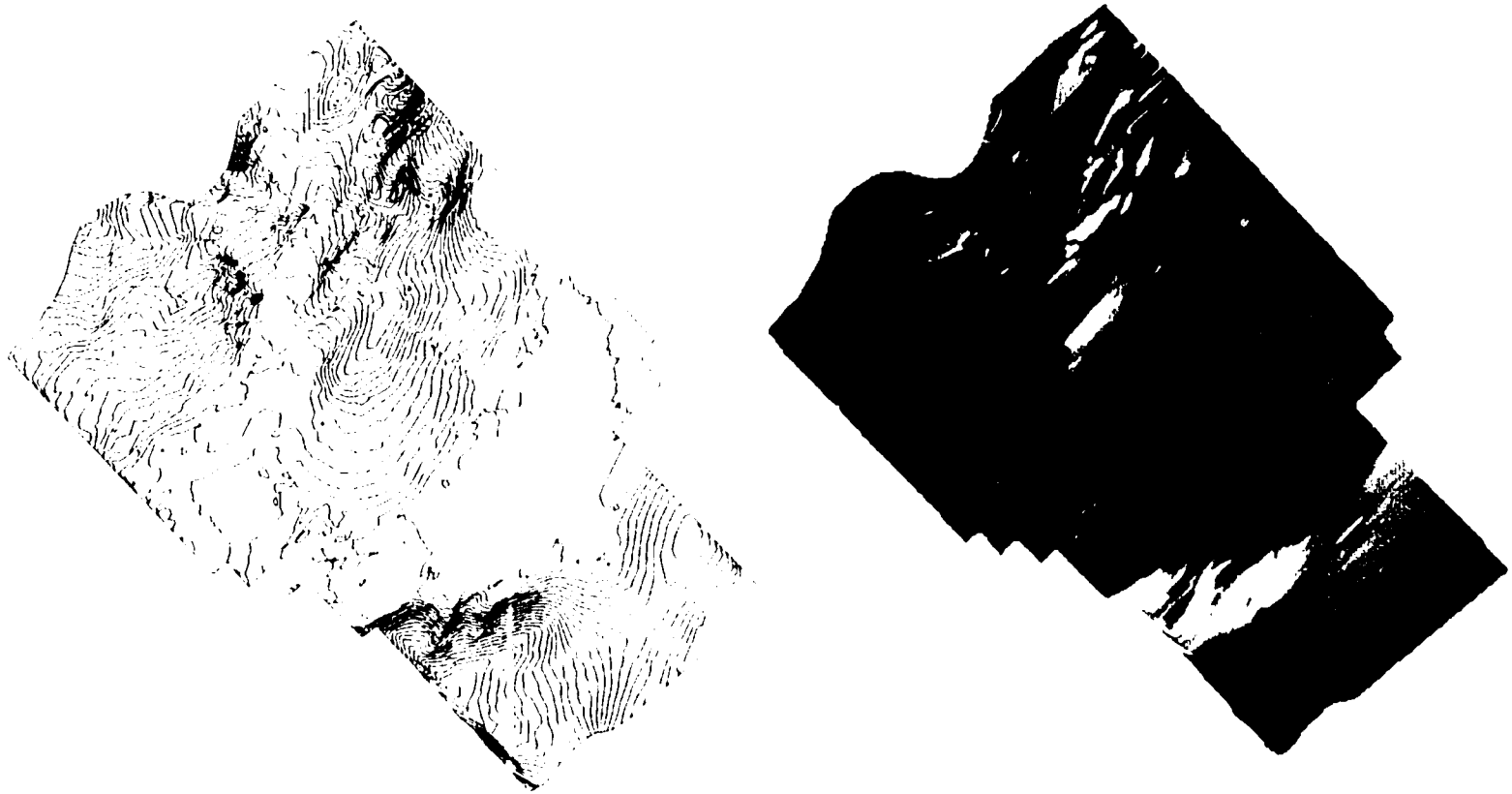
(a) (b)

Fig. 4.8: Interpolation result with minimum curvature spline on the traverse data set.
(a) Contours derived from the interpolated DEM; (b) Hill-shaded image of the interpolated DEM.

extracted from the selected sample points is limited. Furthermore, because the selected sample points are predominantly concentrated in one direction rather than uniformly or evenly distributed around the grid node being estimated, the local interpolation surface is not well determined. As shown in Fig. 4.7 and Fig. 4.8, when Kriging and minimum curvature spline algorithms implemented in ARC/INFO are applied to the traverse airborne radar data over the Evans Ice Stream and Fowler Peninsula, the variation of the resulting surface is very unstable and even erratic. When TIN-based methods are directly applied to the traverse data, the triangulation is characterized by long skinny triangles, generating the unrealistic slopes and artificial strips between traverses (Fig.4.9).

To overcome the drawbacks of the existing interpolation algorithms, a new approach is designed for interpolating the traverse data. This approach involves three steps. First, an algorithm is developed to reduce and filter traverse data along flight tracks. The algorithm is based on a super-block based 2-D searching scheme, as described in Chapter 3. All data points are first partitioned into an array of square blocks. For each block, only one point, the median of all sample points inside the block, is kept. Due to the selection of the median point, the redundant and noisy data points are filtered out. The data reduction level is controlled by the size of the blocks. The array of square blocks is set with a width of one-fifth of the average distance between flight lines.

Second, a directional search based IDW algorithm was developed. It is used to interpolate the airborne traverse radar data into a coarse grid with a spacing of half the average distance between flight lines. To overcome the problem caused by the linear cluster of selected sample points, quadrant and octant search types are incorporated in our implementation of the IDW algorithm. Quadrant search divides sample points around a



(a) (b)

Fig. 4.9: Interpolation result with TIN linear option on the traverse data set.
(a) Contours derived from the interpolated DEM; (b) Hill-shaded image of the interpolated DEM.

grid node into four sectors, and finds the closest data points in each of the four sectors during interpolation. Octant search divides sample points around a grid node into eight sectors, and selects a specified number of the closest data points in each octant. Data points beyond the specified nearest points in a sector are ignored, even if the data points in another sector are farther from the grid node being interpolated. Quadrant or octant search methods can generate a subset of sample points that are evenly distributed around the grid node in all directions. A better azimuthal distribution of the selected sample points has avoided the directional bias caused by the simple search method and therefore results in a reliable interpolation surface. In the case of airborne radar data of the Antarctic, the quadrant search method was used to select three nearest points from each quadrant. Namely, the IDW algorithm interpolates the elevation value at the nodes of a grid using 12 neighboring points equally selected from each quadrant:

$$Z_{i,j} = \frac{\sum_{p=1}^{p=12} Z_p d_p^{-n}}{\sum_{p=1}^{p=12} d_p^{-n}} \qquad (4.14)$$

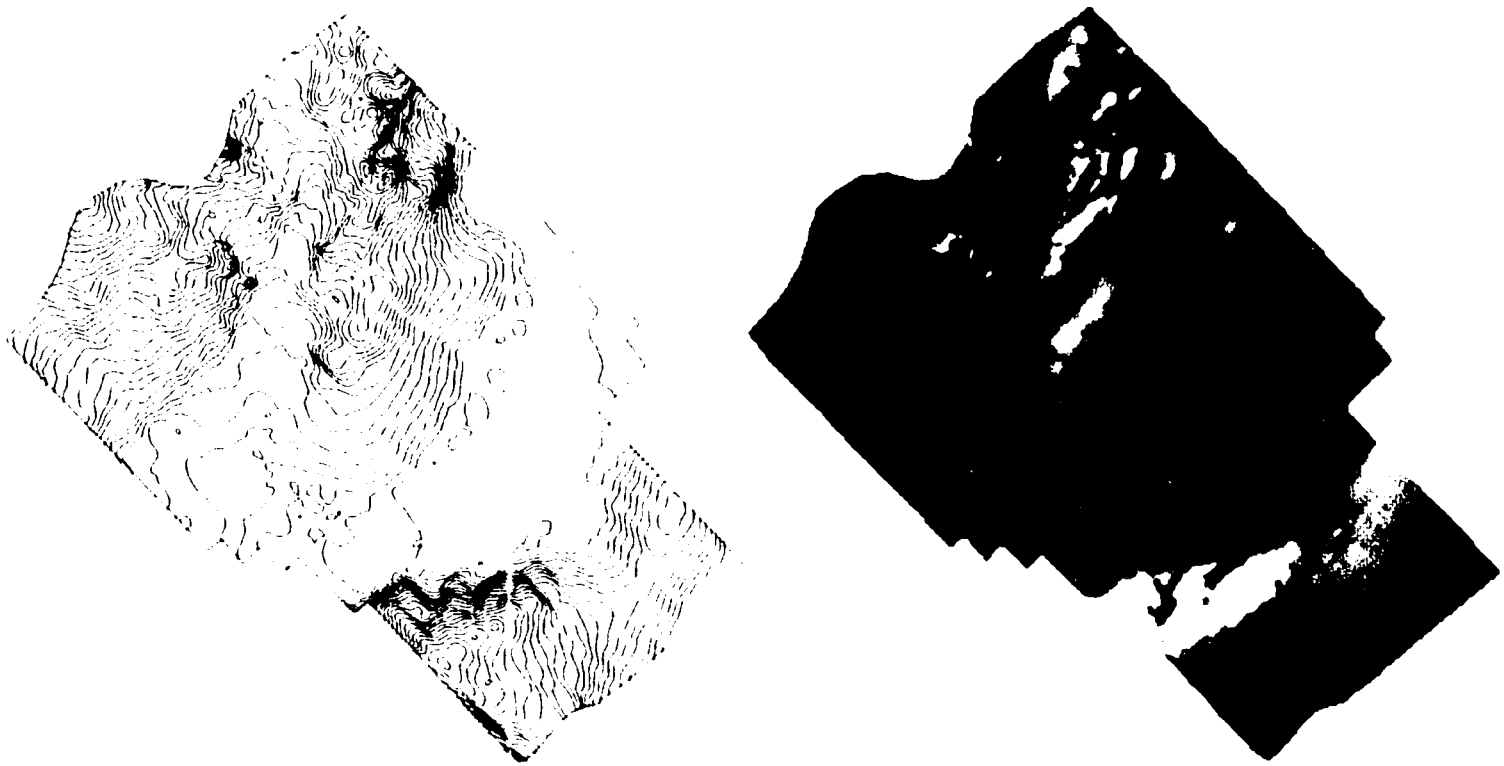
where $Z_{i,j}$ is the computed elevation at the node (i, j) of a grid; Z_p is the elevation at sample point p in the neighborhood; d_p is the distance from the node (i, j) to point p; and n is the "friction of distance". The inverse squared distance weighting, namely, n=2, was used. The sample points in the immediate proximity are assigned higher weights in the interpolation because they are more likely to have similar elevation values.

In the third step, a TIN model was constructed using the reduced radar altimeter data from the first step together with the IDW derived coarse grid points from the second step, and then use the quintic option to interpolate the data into a fine DEM grid.

In our approach, the quadrant search based IDW interpolation serves to stabilize the interpolation result, while the triangulation of the original reduced radar data points and intermediate coarse grid points has the effect of retaining more details of topographical information present in the source data. Therefore, a satisfactory and stable result is achieved (Fig. 4.10).

4.4 Interpolation of contour-based data

Digitized contours are the main data source for DEMs over the rugged mountainous regions of the Antarctic. In comparison with other data types, contour data are the most complex and difficult data type for general-purpose interpolation techniques. Contour data are characterized by the oversampling of information along contour lines and the undersampling between contour lines. Due to the fixed contour interval, the spatial distribution of contour data is strongly heterogeneous. In flat and low relief areas, contours are sparse and smooth, while in the rugged and high relief areas they are very dense and irregular. The shape and pattern of contours implicitly show the specific topological and morphological properties of the terrain, such as the ridges, valleys, hills and pits. Since contours are usually digitized from paper topographical maps, supplemental spot height points and surface structural lines, such as coastlines, lake shorelines, and streams, are often digitized to complement the contour data. To accurately interpolate contour data, the algorithm must deal with the varying data density and exploit the topological and morphological information inferred from the contour pattern or provided by the surface structural lines.



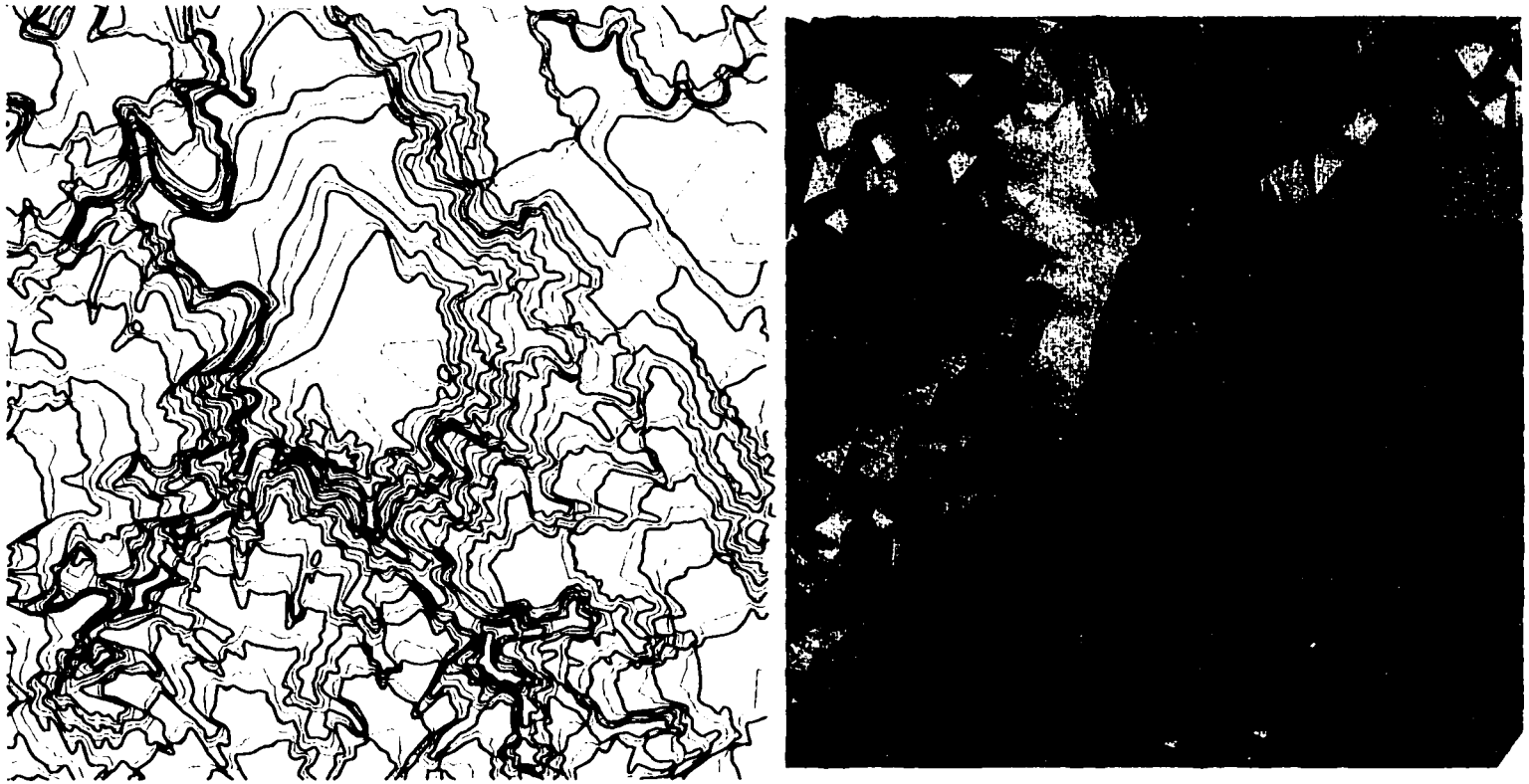
(a) (b)

Fig. 4.10: Interpolation result with the combination of quadrant neighborhood based IDW and TIN quintic on the traverse data set. (a) Contours derived from the interpolated DEM; (b) Hill-shaded image of the interpolated DEM.

We compared three interpolation algorithms that appeared in the literature for interpolating contour-based data at a test site in Antarctica, including the TIN based method (ESRI, 1991a; Auerbach and Schaeben, 1990; Robinson, 1994; Aumann et al., 1991; Carrara et al., 1997), the linear gradient descent interpolation method (Yoeli, 1984; Oswald and Raetzsch, 1984; Eklundh and Martensson, 1995), and the TOPOGRID-based method (Hutchinson, 1988; Hutchinson, 1989; ESRI, 1991b). Among them, TOPOGRID proved the most effective.

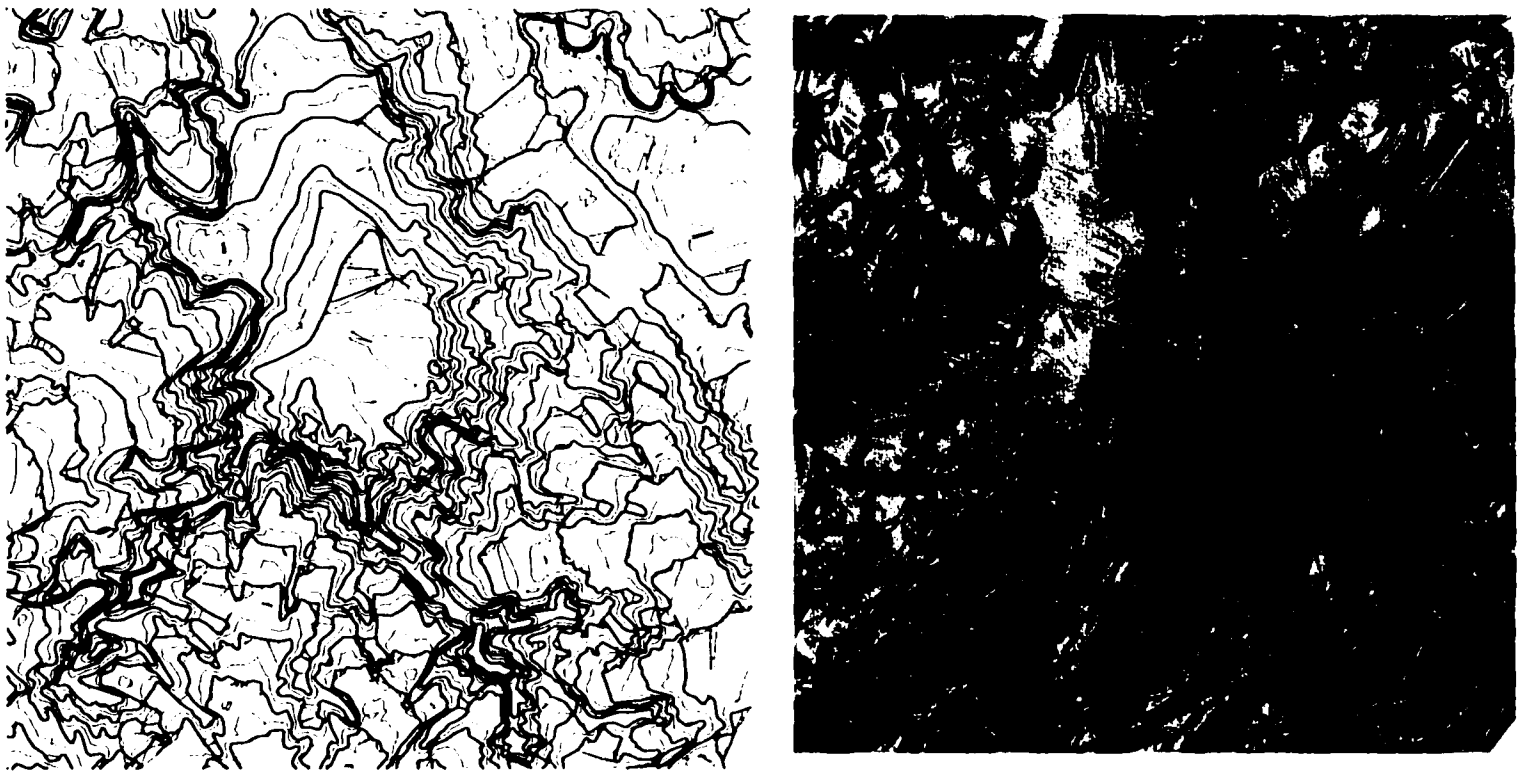
The TIN-based method implemented in ARC/INFO can incorporate contour lines and external surface breaklines as input data to interpolate an elevation grid, but the topological and morphological information implied by contour shape and pattern are ignored. This algorithm has two major drawbacks. First, long skinny triangles emerge in the areas where the contours are widely spaced. As a result, the interpolated values are less reliable and surface slope changes abruptly when crossing over the edge separating two adjacent spiky triangles. Secondly, invalid flat triangles occur where the contour lines have sharp bends, which are normally associated with ridges or valleys. When sample points on the same contour line are closer than the nearest sample points on an adjacent contour line, the Delaunay triangulation process produces flat triangles whose three corner points have the same elevation value. Due to erroneous flat triangles, the fine details of ridges and valleys evident in contours are distorted or lost, and many significant terrace-like artifacts are generated (Fig. 4.11 and Fig. 4.12).

The linear gradient descent interpolation method (Yoeli, 1984;Wood and Fisher, 1993) is similar to what human beings do when visually interpreting a contour map. The algorithm first rasterizes vector contours by assigning the grid cells that a contour line



(a) (b)

Fig. 4.11: Interpolation result with TIN linear option on the contour data. (a) Contours (thin lines) with 100m interval derived from the interpolated DEM are superimposed on the top of the original contours (thick lines) with 200m interval. (b) Hill-shaded image of the interpolated DEM.



(a) (b)

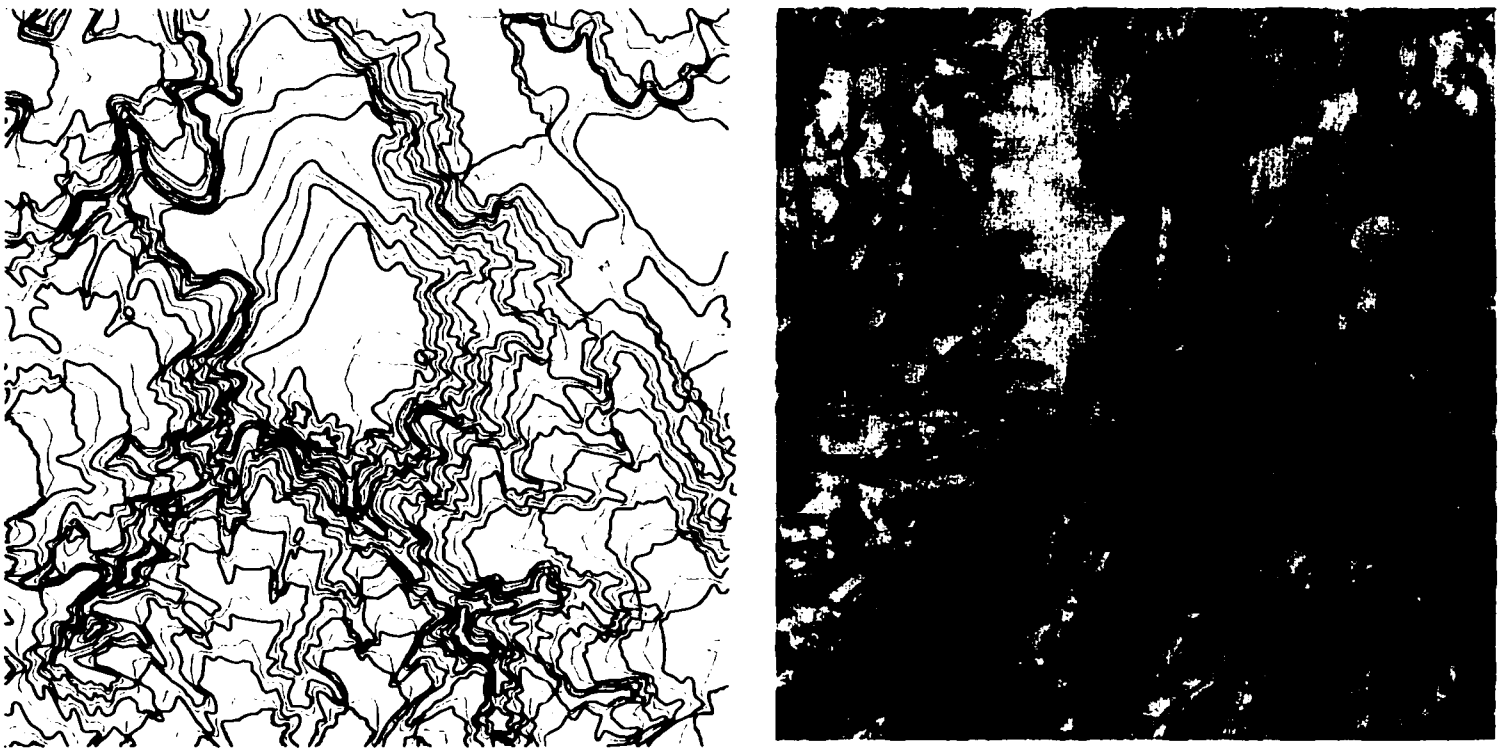
Fig. 4.12: Interpolation result with TIN quintic option on the contour data. (a) Contours (thin lines) with 100m interval derived from the interpolated DEM are superimposed on the top of the original source contours (thick lines) with 200m interval. (b) Hill-shaded image of the interpolated DEM.

cross over with the elevation value of the contour. The values of the remaining cells are interpolated in a flood-filling fashion. For each grid cell between contour lines, the program searches along the directions parallel to the grid rows (E-W), columns (N-S) and two diagonals (NW-SE and NE-SW), and the steepest path is determined by comparing the slopes of these directions. Then, the value of the grid cell is determined by a linear interpolation down the steepest path between two adjacent uphill and downhill contours:

$$Z = \frac{d_{up} Z_{dn} + d_{dn} Z_{up}}{d_{up} + d_{dn}} \tag{4.15}$$

where Z is the interpolated value; d_{up} is the distance from the grid cell being interpolated to the nearest upslope contour; d_{dn} is the distance from the grid cell being interpolated to the nearest downslope contour; and Z_{up} and Z_{dn} are the elevation values of two bounding contour lines. This algorithm gives reasonable interpolation results between contours on the sloping sides of the hills and valleys, and is also able to keep the fine surface structures related to ridges and valleys present in contour data. The major problem with this algorithm is that it is unable to extrapolate the elevation values on the top of hills and bottom of basins. The grid cells bounded by the innermost single contour line are given the same elevation value as the bounding contours, producing unrealistic flat hill tops and basin bottoms (Fig 4.13). In other words, this algorithm truncates the natural convex or concave surfaces associated with local maxima and minima. In addition, this algorithm is computationally very expensive.

The TOPOGRID method implemented in ARC/INFO is based on Hutchinson's (1988; 1989) sophisticated iterative finite difference algorithm. This method uses a nested multi-resolution computation structure, starting with an initial coarse grid and



(a) (b)

Fig. 4.13: Interpolation result with linear gradient descent on the contour data. (a) Contours (thin lines) with 100m interval derived from the interpolated DEM are superimposed on the top of the original source contours (thick lines) with 200m interval.

successively halving the grid spacing until the final specified grid resolution is obtained. At each resolution, the starting values for each successive finer grid are bi-linearly interpolated from the preceeding coarser grid. The data points are allocated to the nearest grid cell and values at grid cells not occupied by data points are calculated by Gauss-Seidel iteration with over-relaxation, subject to a rotation invariant roughness penalty and ordered chain constraints. The roughness penalty is defined by the second order partial derivatives of the fitted function:

$$J(f) = \iint (f_{xx}^2 + 2f_{xy} + f_{yy}^2)dxdy \qquad (4.16)$$

where f_{xx}, f_{xy}, and f_{yy} are second derivative terms. The minimization of the roughness function *J(f)* leads to the minimum curvature interpolation of a thin plate spline. The tendency of the minimum curvature of thin plate spline to smooth out the ridges and valleys is corrected by enforcing linear interpolation along all ridge lines and stream lines, which are automatically derived from points with maximum local curvature on contour lines. In other words, the topographical and morphological information inferred from the shape and pattern of contours are used in TOPOGRID method. The replacement of the minimum curvature constraint with linear gradient descent interpolation allows the fitted surface to have sharp changes in slope along ridge and streamlines. The other feature of this algorithm is that an automatic drainage enforcement procedure is used to remove spurious sinks or pits in the fitted elevation surface.

One major drawback with TOPOGRID method is fictitious overshoot and undershoot in the areas where the contour density has a sharp change. For example, spurious undershoots are generated in the contour sparse areas with very strong relief

nearby, notably in the low slope floors of U-shaped glacial valleys flanked by steep glacier shoulders (Fig. 4.14). This problem has been also observed and reported by other research workers (Bliss and Olsen, 1996). The tendency of minimum curvature interpolation to maintain regional trends is the cause of the problem.

A modified two-stage TOPOGRID interpolation method was designed to overcome the undershooting problem through the adjustment of contour density. At the first stage, the contour lines in low slope regions are densified. A very coarse grid is first interpolated from contour lines using the TOPOGRID algorithm. Then, the surface slope is calculated, and continuous low slope regions are delineated using a region-growing algorithm. The density of contour lines is increased by four times in the low slope regions by converting the coarse grid of the selected low slope regions into contour lines using one fourth of the original contour interval. At the second stage, all available input data are integrated to create a final high-resolution DEM grid. The program is able to incorporate the original contours, the densified contours in the low-slope regions, spot elevation points, grounding line and coastlines. This two-stage procedure is automated in the GIS environment through an ARC Macro Language (AML) program. It combines a number of ARC/INFO internal functions, including TOPOGRID interpolation, slope calculation, region expansion and growing, and contouring.

Streamlines and lake shorelines in coastal mountains contained in the ADD database are also used to constrain the interpolation where they are available. The elevation of coastlines is set to zero. The grounding lines, indicating the locations where terrain surface changes from relative flat ice shelves to the sloped coastal margins of the ice sheet, are used as important structural lines in interpolation. They are first divided into

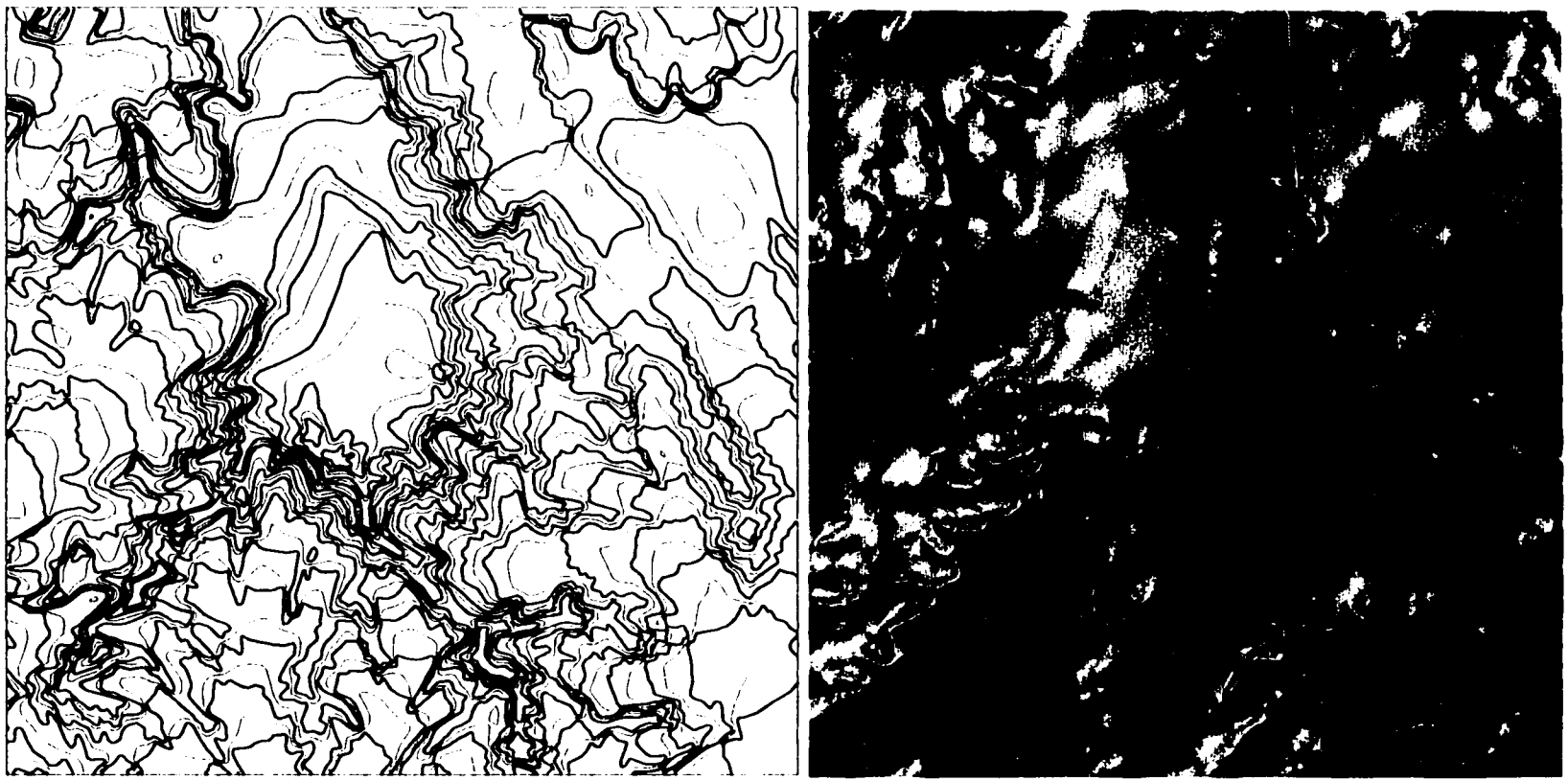
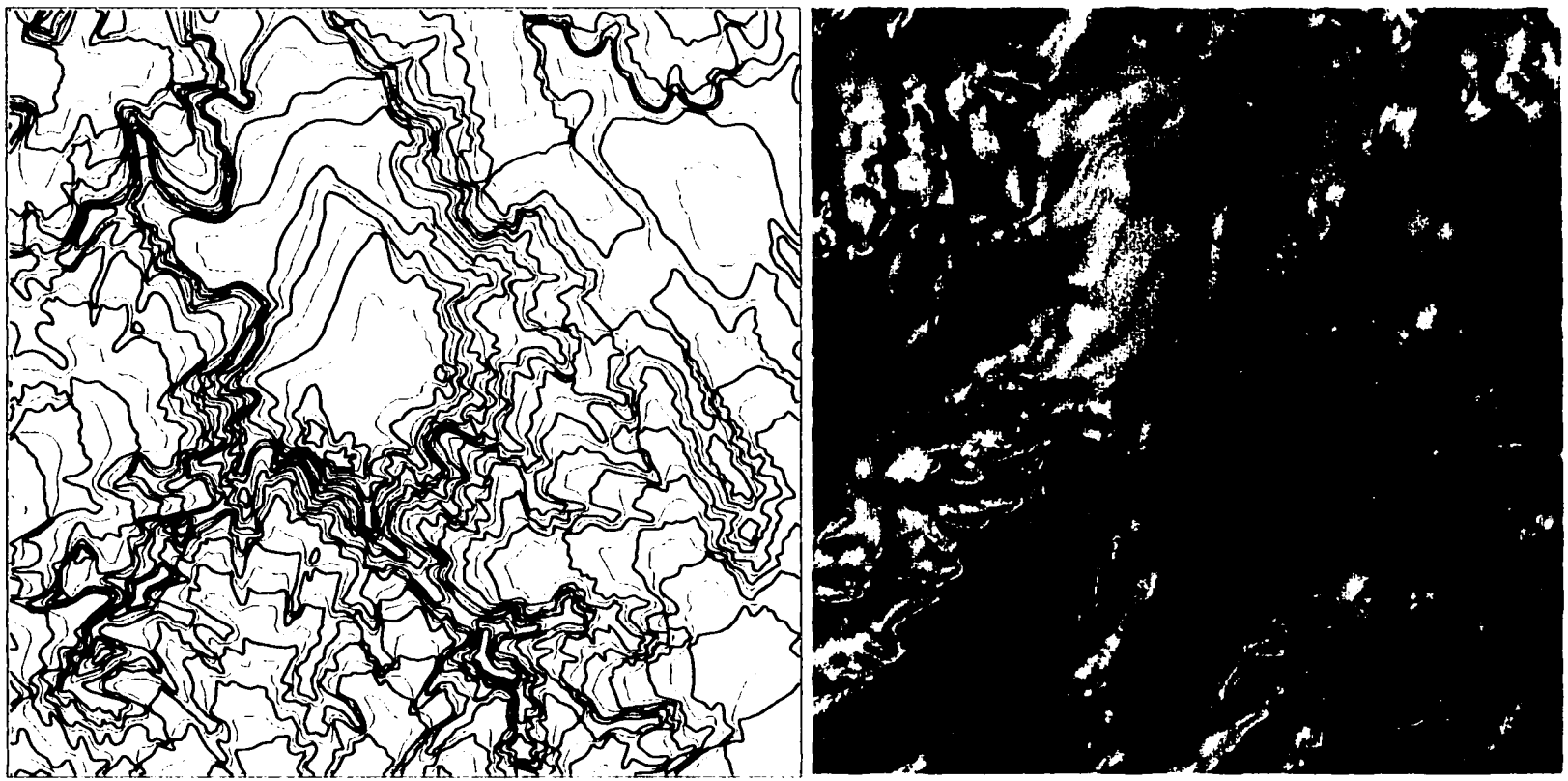


Fig. 4.14: Interpolation result with TOPOGRID on the contour data. (a) Contours (thin lines) with 100m interval derived from the interpolated DEM are superimposed on the top of the original source contours (thick lines) with 200m interval.

small segments, and then the segments are matched with the nearest satellite radar altimeter data points in the ice shelves. The elevation value of the nearest point is transferred to the corresponding grounding line segment. The above preprocessing steps for grounding lines and coastlines are also performed in the GIS environment in an automated fashion through an AML program, which combines a number of ARC/INFO internal functions in line densification, nearest point searching, point and line coverage matching, and PAT and AAT attribute table manipulations.

This modified two-stage TOPOGRID method leads to a geomorphologically correct DEM with several desirable properties (Fig. 4.15). First, the detailed and fine structures of the topographic surface contained in the original source data are preserved and enhanced in the resultant DEM. This is because the replacement of the minimum curvature constraint with linear gradient descent interpolation in the TOPOGRID algorithm allows the fitted surface to have sharp changes in slope along surface structural lines, and because a significant amount of surface morphologic information was incorporated in the interpolation, including ridge and stream lines automatically derived from contours, grounding lines, and surface peaks in spot height data sets. Second, owing to the surface smoothness constraint imposed by the minimum curvature of thin plate spline, the resulting surfaces are continuous and differentiable, and hence visually smooth. Third, the densification of contour lines in low-relief regions alleviates or eliminates artificial pits. The reliability and effectiveness of this procedure is shown by comparison of the interpolated contours with original contours in Fig. 4.15. In the figure, the interpolated contours are twice as dense as the original ones. Those with the same elevation labels overlap the original contours and are indistinguishable from the original



(a) (b)

Fig. 4.15: Interpolation result with modified two-stage TOPOGRID procedure on the contour data. (a) Contours (thin lines) with 100m interval derived from the interpolated DEM are superimposed on the top of the original source contours (thick lines) with 200m interval. (b) Hill-shaded image of the interpolated DEM.

ones, while those with the intermediate elevation labels are consistent with the original ones in shape and pattern.

4.5 Determination of optimal interpolation interval

The grid spacing (cell size) is an important parameter that we have to specify during interpolation. Technically, we can interpolate the source data into a grid with an arbitrarily small horizontal spacing. However, if the grid spacing is too small, the output data volume will expand rapidly. Halving the grid spacing will lead to an increase in data volume by four times. This will in turn increase the storage and subsequent computational costs. In addition, the use of a too small grid spacing will virtually reduce the relative density of sample data points and tends to give rise to interpolation artifacts. On the other hand, if the post spacing is too big, we will lose information contained in the original topographical source data. Therefore, we need to pursue an optimal or appropriate grid spacing at which we can minimize the data volume without a significant loss of information.

For direct data acquisition, the data sampling rate of measuring sensors should be determined by the maximum rate of terrain variation, namely, the minimum topographical feature to be resolved. According to the Whittaker-Shannon sampling theorem (Weaver, 1983), the sample spacing must be smaller than half the wavelength of the highest frequency component of surface topographical variation. Otherwise, the frequency pattern will be distorted due to the unwelcome aliasing effects. For a rugged and mountainous terrain, a small sampling interval is required to obtain an accurate representation of the surface details. For a low-relief terrain, a large sampling interval

may be sufficient. For the interpolation of collected topographical source data into a grid, the grid spacing should be chosen according to the spatial resolution of source data, namely, the amount of topographical information content contained in the source data. The selection of a grid spacing should ensure that the inherent topographical details present in the source data could be faithfully transferred into the elevation grid without introducing artifacts.

In the case of point-type data, the density and pattern of the data reflect the level of topographic details. To keep the topographical information, the grid spacing used for interpolation is often set about two to five times smaller than the average spacing of source data points in practice (Ackermann, 1996). For the preprocessed satellite radar altimeter data, a spacing of 1 km was used for interpolation. This is smaller than half the average distance between data points (2.5 km). It should be noted here that the use of a smaller grid spacing cannot increase the original spatial resolution of the source data, but it does prevent losing the details contained in the source data.

In the case of contour data, the topographic information content is determined by the quality of the original measurements, the map scale and contour interval. Since the horizontal distance between contours varies from place to place, depending on the surface slope and curvatures, there is no straightforward way to determine an optimal grid spacing for interpolation. We performed two analyses on a test site around Mt. Markham in the Transantarctic Mountains for the determination of an appropriate grid spacing for interpolation.

First, a Fourier analysis was conducted to diagnose the information content of contour data. The size of the test site is 51.2 km by 51.2 km. A fine DEM grid

(2048x2048) with 25 m grid spacing is first created by the TOPOGRID method based on 1:250,000 scale contour data that is the typical scale for the Transantarctic Mountains and Antarctic Peninsula. The mean elevation of this DEM grid is 1710 m, and the standard deviation of elevation is 840 m. This fine grid contains almost all the topographical information present in the source data, which is verified by the close matching and consistency between the derived contours and original contours. Then, the FFT (Fast Fourier Transformation) is performed in the frequency domain. The surface power (variance) at each frequency is calculated, and the radial power spectrum is created (Fig. 4.16). Higher frequency terms relate to the finer details of the topography, whereas the lower frequencies contribute more to the overall shape. According to the Parseval-Rayleigh theorem, the sum of the variance of surface variation computed in the space domain equals the total power of the surface (squared amplitudes) calculated in the frequency domain (Weaver, 1983). At the cut-off frequency, k, the lost information (variance) can be calculated by (Frederiksen, 1981; Balce, 1987):

$$\sigma^2_{lost} = \sum_{k_1=k_2=k_{cut-off}}^{N/2} P(k_1,k_2) \tag{4.17}$$

where $P(k_1,k_2)$ is the power density at frequency (k_1,k_2).

If we maintain 99.9% information (variance) transfer from the source data to the reconstructed grid elevation model, the corresponding cut-off frequency ($k_{cut-off}$) is 160 cycles per 51.2 km for the test data (Fig. 4.16). The optimal post spacing $\Delta X_{optimal}$ can be then calculated as:

$$\Delta X_{optimal} = \frac{N}{2k_{cut-off}} * \Delta X_{original} \tag{4.18}$$

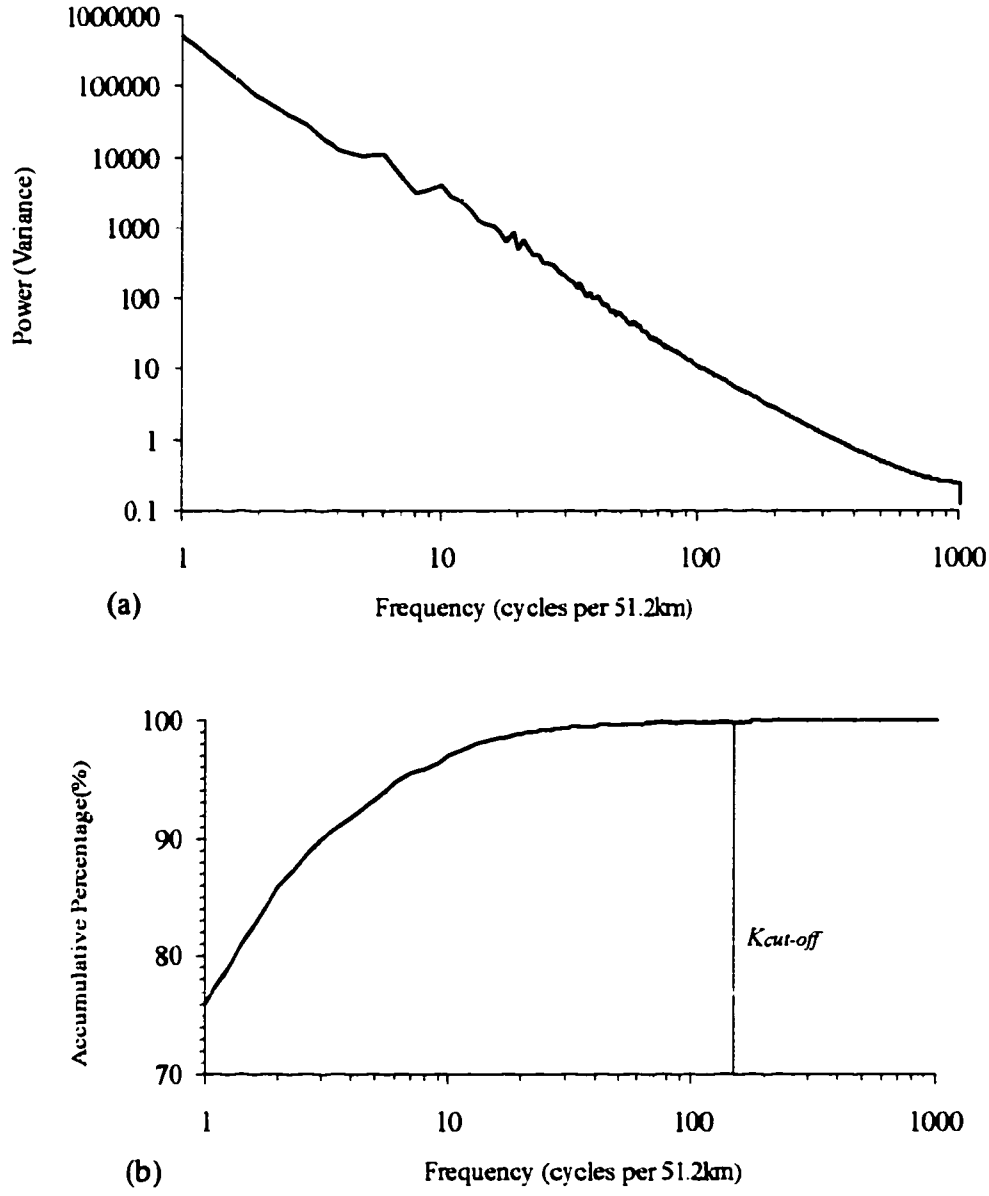


Fig. 4.16: Radial power spectrum of mountainous terrain.
(a) Power (variance) change with frequency, and unit is Cycles/51.2 km; Y-axis: power (variance) in logarithmic scale.
(b) Accumulative percentage of power. X-axis: frequency in logarithmic scale; Y-axis: accumulative percentage.

where $\Delta X_{original}$ is the original sampling spacing (25 m), and N is the sampling points (2048) along each side. From the test data, we obtained the optimal grid spacing of 160 m. The loss of 0.1% information (variance) result in an elevation uncertainty of about 26.5 m (standard deviation).

Second, a multi-resolution analysis was carried out to analyze the relationship between interpolation error and grid spacings. For the same test data, the contour data were interpolated respectively into 50 m, 100 m, 200 m, 400 m, 800 m, 1600 m and 3200 m grids using the TOPOGRID algorithm. Then, the differences between these successively coarser grids and 25 m grid, which is used as the base reference grid, are calculated. The average height difference is used as a measure to quantify the interpolation errors caused by using a larger grid spacing. As shown in Fig. 4.17, the interpolation error is found to increase linearly with the increase of the grid spacing. The 200 m grid spacing corresponds to the average height error of 16.4 m.

Based on the above analyses, we used 200 m as the interpolation grid spacing for mountainous areas where 1:250,000 scale contour data are available. It should be noted that the test area is one of the most rugged areas in the Antarctic, where the contour density is very high. For the sloped coastal area, the terrain surface is relatively smooth and contours are relatively sparse and regular, so a 400 m grid spacing is used instead. This is again verified by comparing the derived contours with original contours.

4.6 Summary

A set of alternative spatial interpolation algorithms are incorporated in the IGIAS based on the ARC/INFO internal functions and C language programming. The experiments

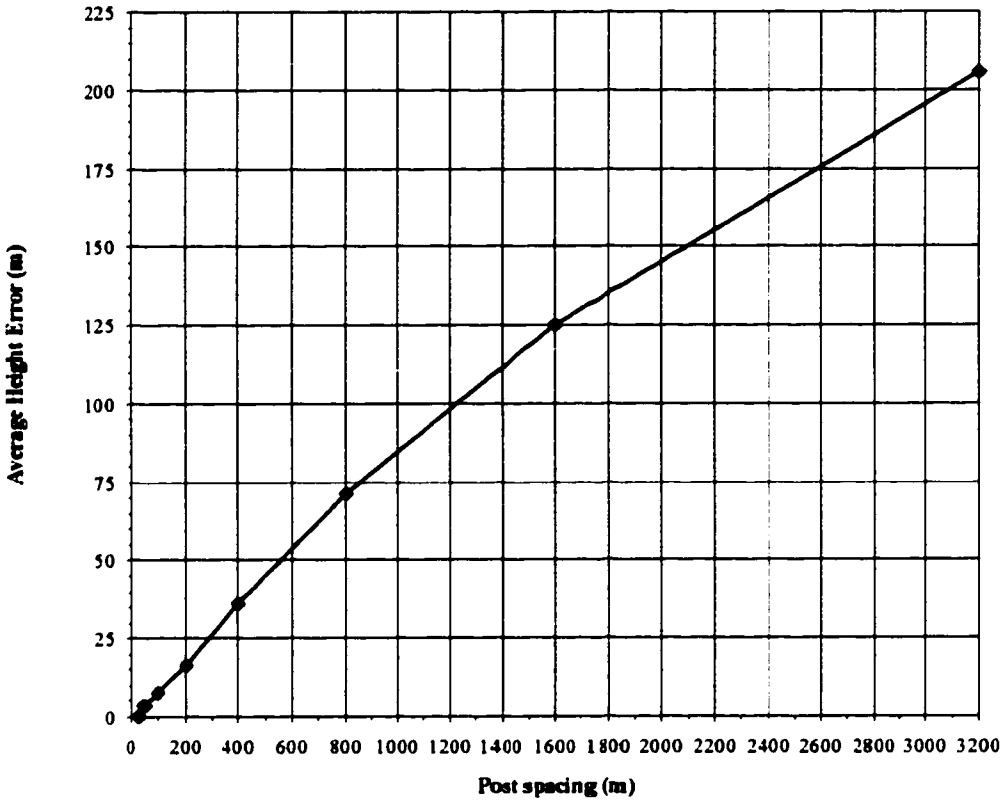


Fig. 4.17: Changes in average height error with interpolation post spacing

with these algorithms were conducted in terms of different types of source data, and their performances were analyzed. The available ARC/INFO interpolation methods only give the satisfactory result for the evenly distributed data set, but none of them can produce acceptable results for traverse and contour data. Therefore, a new interpolation approach was developed for traverse data by combining the median filtering, directional search based IDW and TIN quintic interpolations. Also, a modified two-stage TOPOGRID interpolation procedure was designed for contour-based data by combining a number of ARC/INFO internal functions through an AML program. In addition, the power spectrum analysis and multi-resolution analysis were carried out to determine an appropriate grid spacing (cell size) for interpolation. Through the development and use of the IGIAS software package, we demonstrated that the GIS technology has a strong capability of retrieving and incorporating various topographical data, point-type or line-type, direct measurements or surface structural lines, into the interpolation process. We also point out that the GIS technology provides an environment for integrating different interpolation algorithms under a common interface and a feedback mechanism for adjusting interpolation parameters and for identifying a suitable interpolation algorithm through the feedback circle-interpolation, visualization, verification, adjustment, and then interpolation.

CHAPTER 5

DATA INTEGRATION AND MERGING

5.1 Conceptual framework for data integration

As described in Chapter 3, each topographical data acquisition technique has a limited data sampling density and an inherent measuring accuracy, and may function incorrectly under some conditions due to specific types of noise and errors. Topographical data extracted from different techniques and in different periods of time often cover different parts of the ground surface. This is the case with the topographical database that we compiled for the Antarctic. The components of the database, including satellite radar altimeter data, airborne radar data, digitized contours, supplemental spot height points, various surface structural lines, ground survey data and GPS data, have varying accuracy, spatial resolution and ground coverage. To make a continental-scale digital elevation model of best possible quality, we need to integrate all available data components from a variety of sources in an optimal way such that the comparative advantages of each source can be fully exploited. The synergistic combination of topographical measurements from various sources can generate a digital elevation model that is much more reliable and accurate than what a single source could create alone.

A variety of methods and techniques can be used for integrating multiple data sources. The design of integration mechanisms is related to not only the characteristics of each

source, but also the relationships between different sources. A sensible analysis of their relationships is the prerequisite for formulating a concrete framework for integrating topographical data under different circumstances.

This research classifies the relationships between different topographical data sets into two types: complementary and competitive. This classification scheme is adapted from Durrant-Whyte's (1988) work in multiple sensor fusion. Two topographical data sets may be spatially or functionally complementary. When two data sets cover different geographic regions, they are spatially complementary. If two data sets cover the same area but each of them has its own relative advantages in some aspects, they are functionally complementary. If two data sets cover the same area but have different vertical accuracy and spatial resolution, they are competitive as to which one will be selected. For spatially complementary data, the objective of data integration is to expand the data coverage. The primary requirement for an integration technique is to merge separate data sets together into a larger consistent data set. For functionally complementary data, the data integration aims to take the comparative advantages of each data set and fuses them into a single data set of higher quality. For competitive data, the data integration involves selecting a better data set based on the comparison and evaluation. The full understanding of the characteristics of each data set is the key to formulate the data selection criteria.

The integration of topographical data may be performed at different stages during the DEM generation process. At the data acquisition and collection stage, the information from one topographical source can guide and facilitate the extraction of elevation measurements from other raw data sources. This type of low-level data fusion can

improve the quality of the extracted surface measurements. At the interpolation stage, various topographical measurements that are spatially and/or functionally complementary can be used together as input to model terrain surface and produce elevation grids. At the output stage, the individual DEM blocks separately interpolated from topographical source data can be coalesced into a DEM mosaic.

After a series of preprocessing as described in Chapter 3, the diverse topographical data sources of the Antarctic have been cast into a coherent and compatible ARC/INFO format and structure. They share a common planimetric and vertical reference systems, polar stereographic projection and OSU91A geoid. With consistent and compatible multiple data sets, a wide range of data integration operations can be conducted.

5.2 Integration of complementary source data

When the overlapped data sources do not depend on each other and have comparable accuracy, they can be used together to densify the input data and/or impose surface topological and morphological constraints during the interpolation. This type of synergistic fusion of different data sources can generate the surface elevation model that is much more than just the sum of the topographical information provided separately by each source. There are a number of data layers in the ADD, including contours, spot height points, grounding lines, coastlines, and a limited number of streamlines and lake shorelines. They are functionally complementary of each other in defining terrain surface.

On the ice shelves, the only available topographical data are satellite radar altimeter data. For the portion of Ross Ice Shelf inside the 81.4°S latitude circle, the only available topographic data is the airborne and station radar sounding data acquired by the RIGGS

project. The ice sheet inside the 81.4°S latitude circle is only covered by small-scale topographical maps of the ADD. Over the surrounding ocean, the surface topography is measured by the ERS-1 ocean-mode radar altimeter data relative to the WGS84 ellipsoid. The data sets listed above are unique and do not overlap with other data sets, therefore, they are spatially complementary for the purpose of obtaining a complete coverage of the Antarctic topography.

The synergistic fusion of various cartographic data was performed in the mountainous and sloped coastal margins of the Antarctic in the GIS environment. During the interpolation, contour data, spot elevation points, coastlines, grounding lines, streamlines, lake shore lines, and ridge and valley lines automatically derived from contours were integrated. Fig. 5.1 illustrates a typical input data pattern in the Antarctic Peninsula where multiple data sources are fused using the modified two-stage TOPOGRID algorithm. The dense contour data provide the basic topographical information and control the overall shape of the terrain surface. Since many of the supplemental spot height points are located in the critical points like mountain peaks and basin pits, they complement the contour data and give much more accurate results in the hill tops and valley bottoms. The coastlines provide not only the boundary elevation value but also the location of the interpolation extent that prevents the extrapolation errors. Grounding lines, streamlines, shorelines, ridgelines and valley breaklines derived from the curvature property of contours provide the topological and morphological constraints on the interpolation process. This serves to avoid the distortion of the fine surface structure and retain the topographical details present in the contour data. The satellite ice-mode radar altimeter

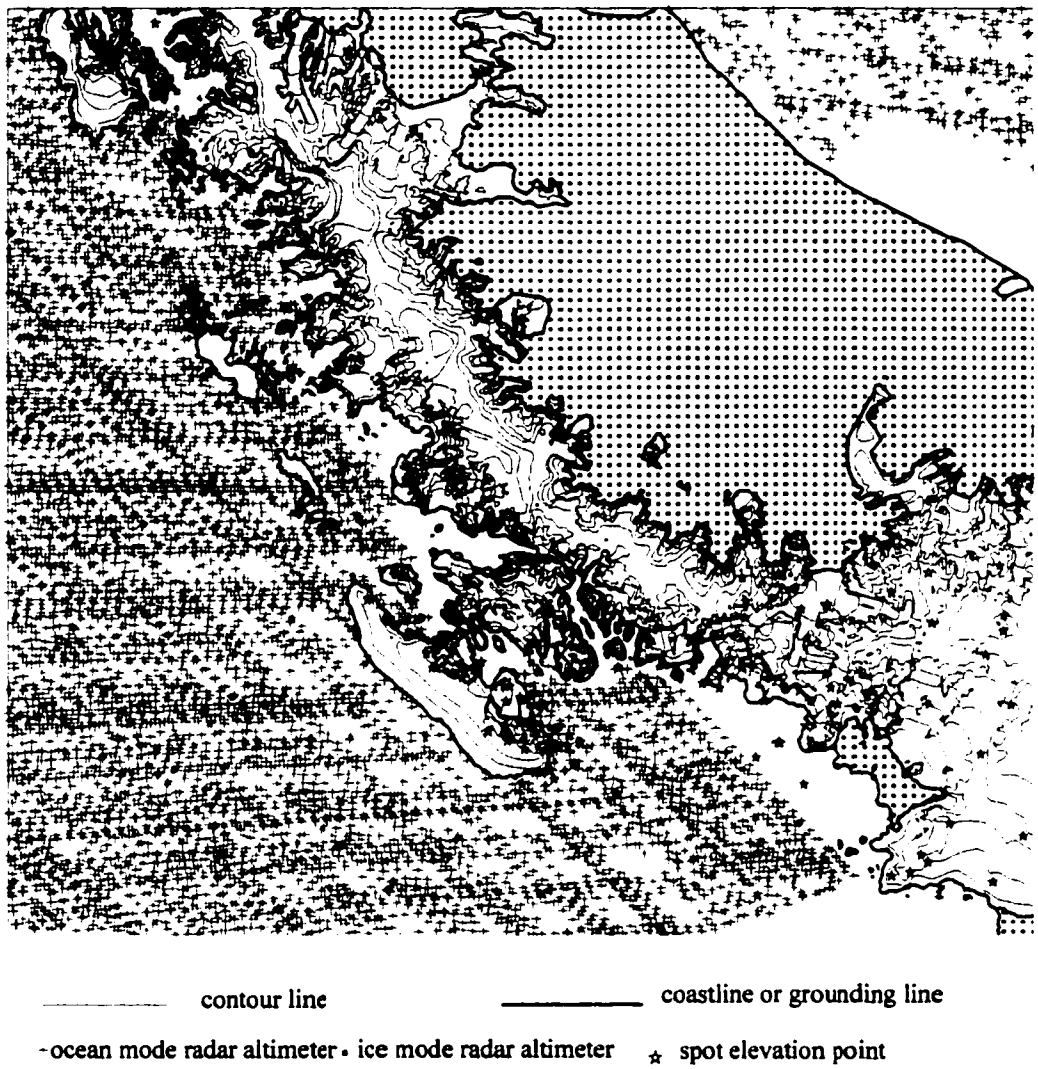


Fig. 5.1: Integration of various input data sources for DEM creation in Antarctic Peninsula

data over the ice shelf and ocean-mode radar altimeter data over the offshore ocean area are spatially complementary to the cartographic data to cover the entire area.

5.3 Selection of competitive source data

Competitive sources provide multiple measurements of surface topography in the same area. If competitive data sets are derived using different data extraction techniques, their spatial resolution and accuracy might be quite different, and some discrepancies and conflicting measurements might arise. If one data source is absolutely superior in terms of density and accuracy, it will be selected as input in the subsequent surface modeling and other overlapping data sets will be excluded.

In the Antarctic topographical database, the cartographic data with varying scale and accuracy cover the entire continent except for ice shelves and surrounding ocean. The ERS-1 satellite radar altimeter data are distributed from the surrounding ocean to the inland ice sheet up to a latitude of 81.4°S. On the ice sheet, satellite radar data overlap with the cartographic data from the ADD. Airborne radar data, large-scale contour data from the USGS and the Australian Antarctic Division, and GPS measurements are available in limited regions, and they are overlapping with the cartographic data from the ADD and/or satellite radar altimeter data.

Based on the quality assessments, the following criteria for the selection of data sources derived from different acquisition techniques were formulated:

1) use GPS data, airborne radar data, and large scale topographic maps wherever they are available;

2) use satellite radar altimeter data if surface slope is less than 0.8°;

3) use the ADD cartographic data for rugged and highly sloped areas; and

4) use satellite radar altimeter data for areas with surface slopes between 0.8° and 1.0°, if other data sources are sparse, or of poor quality, or do not agree with the surface shape suggested by satellite image data.

Fig. 5.2 shows the distribution of the selected source data used in the final DEM. The accuracy of satellite radar altimeter data is strongly correlated with the surface slope and ruggedness. Although the Antarctic Peninsula, most parts of the Transantarctic mountains and sloping coastal margins are covered by satellite radar altimeter data, the measurements are not reliable or seriously in error due to poor tracking and the unpredictable complex shape of the waveform. In addition, the slope correction algorithms tend to produce gross errors because of the incorrect estimates of local surface slope. Cartographic data in the mountainous and coastal regions are typically at the scale of 1:250,000, originally derived from airborne photogrammetry methods. In comparison, they are much more reliable and accurate than satellite radar altimeter data, and hence used as the input source data for computing the DEM in these rugged and highly sloped areas. The Dry Valley region in the Transantarctic Mountains, the Vestfold Hills, the Windmill Islands, and the Larsmann Hills in the East Antarctica are covered by 1:50,000~1:10,000 scale topographical maps that originate from airborne photogrammetric measurements. The scale and spatial resolution are much better than their counterparts in the ADD, and thus used as the sole data source for these areas.

Three sets of ERS-1 radar altimeter data that we collected were independently preprocessed by three different investigators. The comparison between them has revealed data "cliffs" in one data set due to the imperfection of their data merging scheme

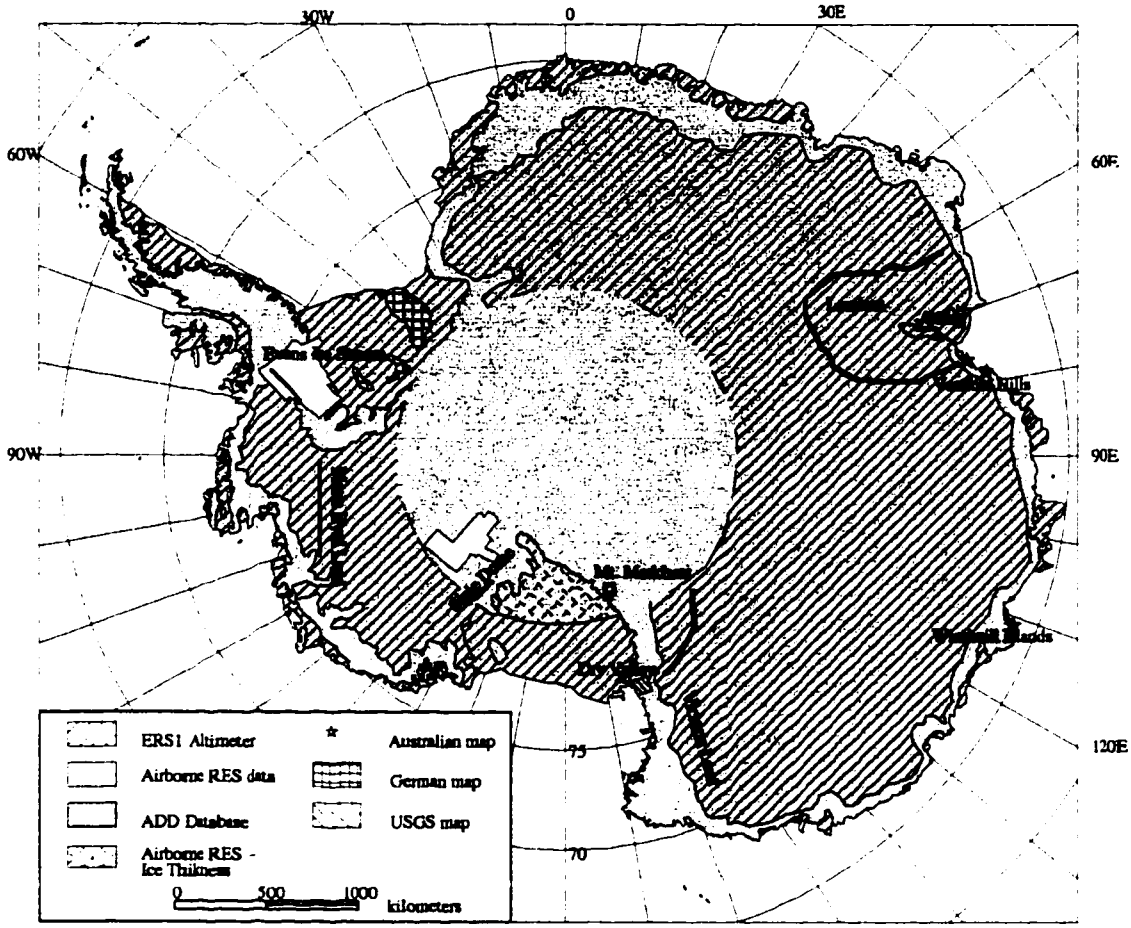


Fig. 5.2: Distribution of input data sources used in the final DEM

and systematic errors of another data set caused by their confusion of the vertical reference system. We reported the problem to the data producer, and the problem has been fixed. Although three data sets cover the same areal extent, the amount of input measurements and the techniques used for waveform retracking and slope correction are different. We selected the data set preprocessed by Zwally et al. (1997). This is because they used more input source data, namely, both two 168-day cycles and two 35-day cycles, and because the final version of their data compared favorably with selected independent data of higher quality.

Conventionally, it is recommended that satellite altimetry can be reliably used in the area where the surface slope is less than half the altimeter beam width, 0.65° in the case of ERS-1 altimeter data (Bamber, 1994; Ekholm, 1996). However, empirical studies show that after a series of error corrections the radar altimeter can provide reliable measurements over terrain surfaces of up to 0.8~1.0° slope (Martin et al., 1983; Zwally et al, 1983). Ekholm (1996) shows that ERS-1 radar altimeter data are more reliable than the digitized topographic maps on surfaces of up to 1.2° slope in Greenland. We took 0.8° as the basic surface slope threshold for the use of satellite radar altimeter data. In some parts of coastal regions, the use of satellite radar altimeter data was expanded to include the surface with the slope between 0.8° and 1° on the conditions that the cartographic data in the ADD are sparse and at small scale, or simply at odds with the shape indicated by the AVHRR image mosaic (Ferrigno et al., 1996) or Radarsat Quicklook image Mosaic (Jezek, 1998). This expansion mainly occurred in East Antarctica, and is supported by the comparison of leveling and GPS traverses with the ADD contour data and satellite radar altimetry data.

Airborne radar sounding data were used over the upstream parts of Ice Streams A, B, and C in the West Antarctica. Its accuracy is estimated to be 4~6 meters, much superior to the contours in the ADD. The GPS navigated airborne radar echo sounding data over the Evans Ice Stream and Fowler Peninsula are more accurate and denser than satellite radar altimeter data as well as the ADD contours, therefore, it is used as the input data for this region.

To facilitate the data selection process, the fundamental GIS operations, such as overlay, spatial search, grid algebra, conditional combination and logical operations have been widely employed. With the view tool of the IGIAS, all the available data layers are conveniently loaded and superimposed layer by layer in the GIS environment for each part of the Antarctic. Through visual inspection, the spatial and functional relationships between different data layers are determined. The coastlines and grounding lines are used as an important condition layer for clipping, inserting or combining diverse topographical data layers to satisfy the selection criteria listed above. For example, the boundary polygons of ice shelves are used to clip the ice-mode satellite radar altimeter data set so as to keep valid measurements inside the marginal ice shelves and drop possible wrong measurements over the rugged and sloped coastal regions. The polygons of coastlines are used to erase possible invalid measurements of satellite ocean-mode radar altimeter data over the continent and ice shelves, with the result that only valid measurements over the ocean are kept. The boundary polygons of several airborne radar data sets are used to insert the DEM grids derived from airborne data into the DEM grids derived from cartographic data or satellite radar altimeter data. In addition, surface slopes are calculated from the DEM grids respectively derived from cartographic data

and ERS-1 satellite radar altimeter data, and then boundary polygons of the regions that have a surface slope less than 0.8~1° are extracted by using grid logical operations and grid-to-polygon conversion. The boundaries of the extracted polygons are used as an important condition layer to support the data selection and merging of cartographic data and satellite radar altimeter data over the transitional zones between the mountainous and highly sloped coastal regions and the relatively flat interior ice sheet.

5.4 Seamlessly merging individual DEM data sets

As a tremendous volume of input and output data is involved in the large-scale DEM generation process, it is not feasible to generate a complete high-resolution DEM in one run. In practice, we have to divide the study area into a set of sub-regions, and generate the DEM block for each sub-region and then combine them together to get a complete DEM mosaic.

The Antarctic was subdivided into seven blocks. For each block, individual topographical source data were combined or trimmed by using the available GIS merging and clipping functions for vector data. To ensure a smooth connection, a certain amount of overlap was set between adjacent blocks. During the interpolation, a significantly smaller interpolation extent than that of the input source data was specified to avoid edge (boundary) effects.

For those DEM blocks that are created from the same type of source data, the merging can be easily performed along the regular boundary of the individual blocks. However, for those DEM blocks that are generated from a different type of source data, an irregular merging line has to be determined. In the case of the Antarctic, the

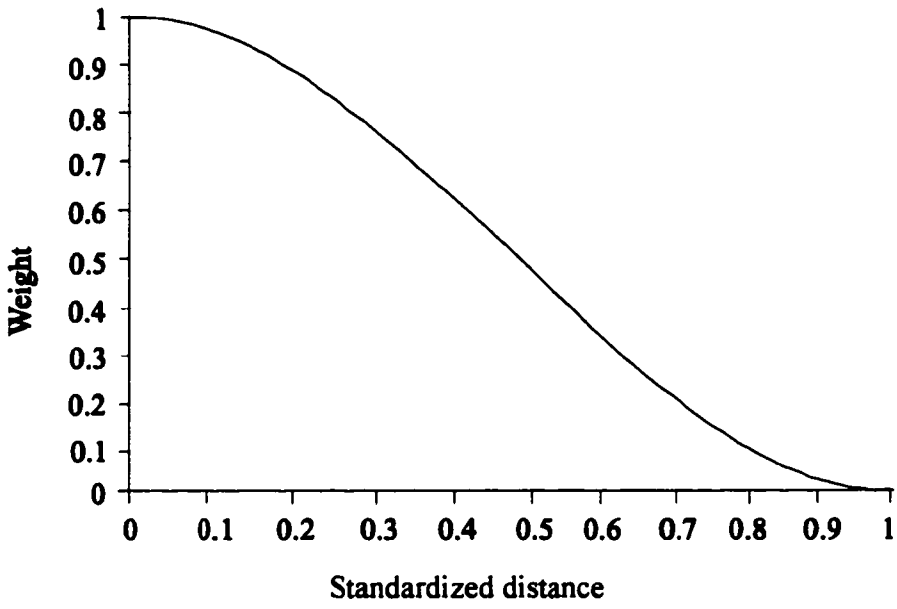
individual DEM blocks of offshore ocean, the peripheral ice shelves and the rugged coastal regions, respectively derived from satellite altimeter data and the ADD cartographic data, were merged along coastlines and grounding lines by using grid logical operations. If a grid cell is located in the ocean, we take the interpolated value from ocean-mode radar altimeter data. If a grid cell is located on an ice shelf bounded by grounding lines and coastlines, we take the interpolated value from ice-mode radar altimeter data. If a grid cell is located on the sloped and rugged coastal region of the continent, we take the interpolated value from cartographic data. The DEM blocks of the rugged coastal regions and the DEM blocks of the flat interior, which are derived from satellite altimeter data, were merged along an irregular buffer zone of about 10~20 km wide. Its location is determined by the surface slope of 0.8~1.0°, the threshold value for the use of satellite radar altimeter data, and modified with reference to the magnitude of the difference between the cartographic data and the satellite radar altimeter data.

To ensure seamless merging, a cubic Hermite function is used to calculate the weights for blending the individual DEM blocks along the irregular buffer zone:

$$w_{sat} = 1 - 3d^2 + 2d^3 \tag{5.1}$$

$$h = w_{sat} h_{sat} + (1 - w_{sat}) h_{ADD} \tag{5.2}$$

where w_{sat} is the weight of satellite radar altimeter data, and d is the distance of an altimeter data point to the buffer line, h is the elevation value after merging, h_{sat} is the elevation value derived from satellite altimeter data, and h_{ADD} is the elevation value derived from the ADD data. As shown in Fig. 5.3, the cubic Hermite function is a S-shaped curve. In the buffer zone, the satellite radar altimeter data have higher weights on the low slope side, while the cartographic data have higher weights on the high slope



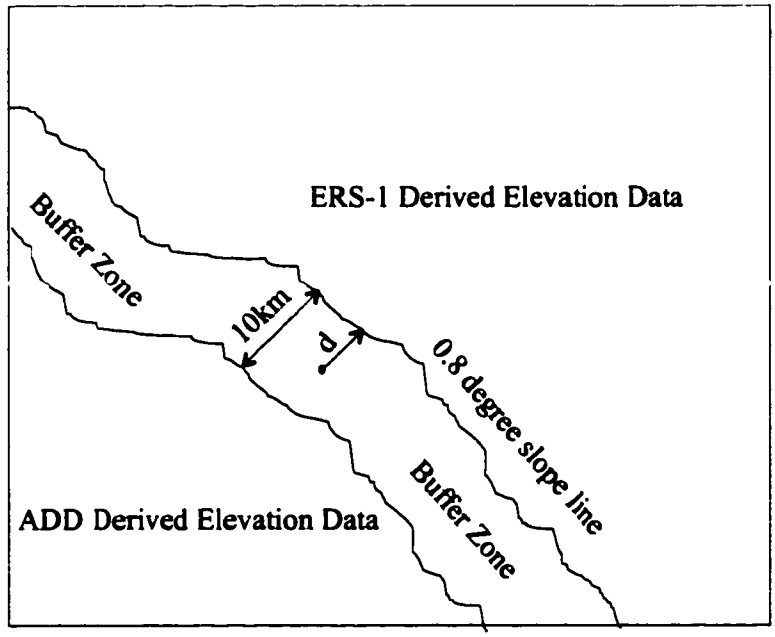


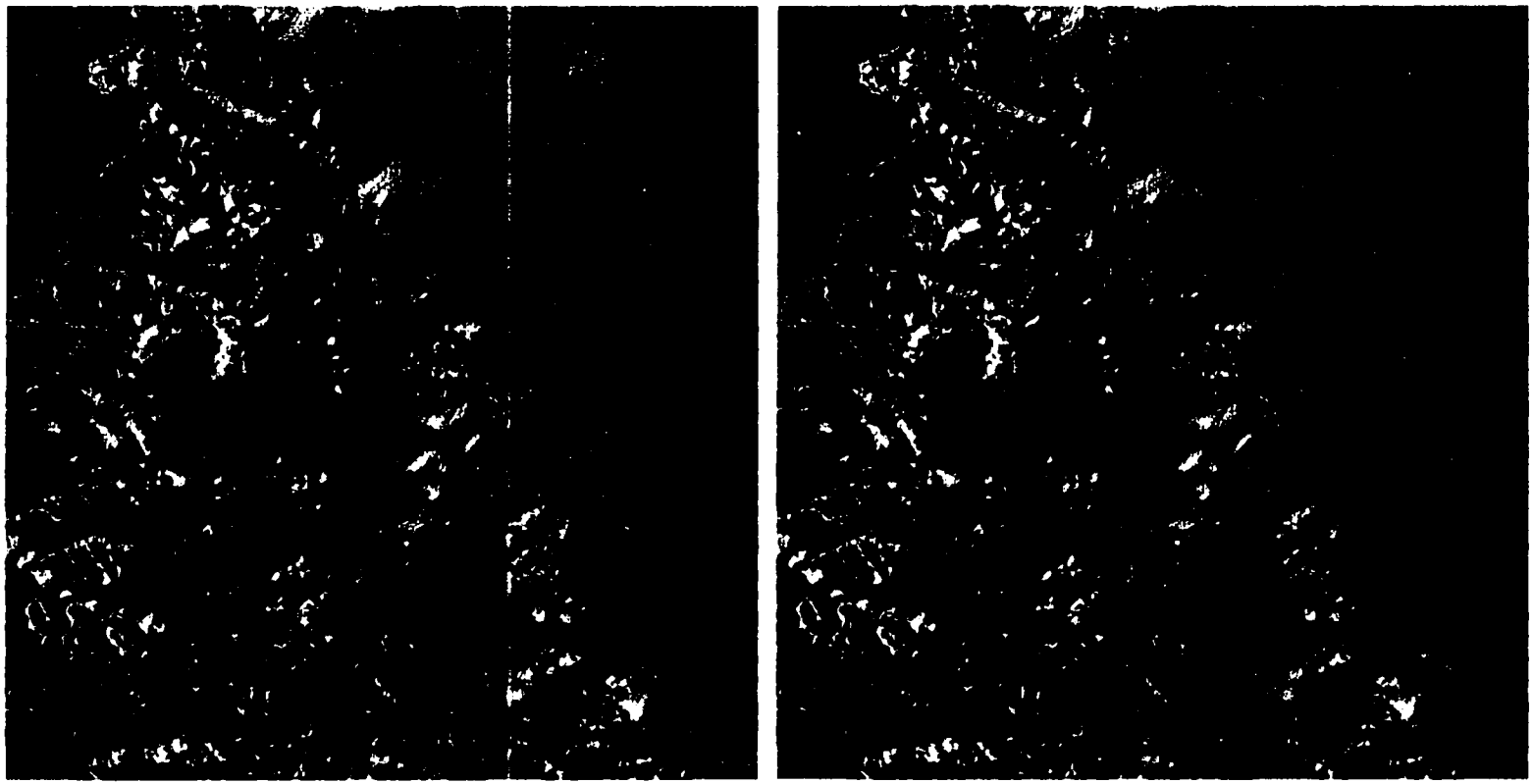
Fig. 5.3: Data merging scheme using cubic Hermite blending function

side. The generation of the weight grid is realized by an AML program by combining a number of available GIS functions, including line and polygon buffering, distance calculation, grid algebraic operations and mathematical transformations.

Fig. 5.4 shows the merging results of using this technique and no seam lines are observed. In contrast, a data step is evident in the merged data set produced by the traditional method.

The grid spacings of individual DEM blocks have been optimally determined for each type of source data during the interpolation stage. As explained in Chapter 4, 200 m, 400 m, and 1 km DEM blocks have been created for the Antarctic, depending on the geographical location and the spatial resolution of the source data. Due to the nature of the raster data structure, the individual DEM blocks must be adjusted to have the same grid spacing and orientation before the merging operation is conducted. The bilinear resampling was used for the adjustment of the grid spacings of these blocks. In comparison with the vector-to-grid interpolation, the bilinear grid-to-grid resampling is computationally much faster and efficient, and also has very little possibility of introducing artifacts.

After the adjustment, the individual DEM blocks are merged at three different levels, resulting in three DEM mosaics respectively with a grid spacing of 200 m, 400 m, and 1 km. For the 200 m DEM mosaic, the individual DEM blocks with 400 m and 1 km grid cells were densified by the resampling process. For the 400 m DEM mosaic, the individual DEM blocks with 200 m spacing were coarsened by the resampling, while the individual DEM blocks with 1 km spacing were densified. It should be noted that the densification of grid cells by the resampling process could not increase the topographical



(a) (b)

Fig. 5.4: Comparison of the cubic Hermite blending function-based merging method with conventional method.
(a)DEM grid merged by the cubic Hermite blending function method; (b)DEM grid merged by conventional method.

information (real resolution), but the coarsening of the DEM blocks by resampling could result in significant loss of the topographical information contained in the original DEM blocks.

In summary, the available fundamental GIS functions in vector data overlay-type operations and grid algebraic, logical and mathematical operations greatly facilitated the data preparation, selection, merging and integration that are required for the generation of a large-scale DEM. The data selection process ensures that the best topographical source data are chosen to model the terrain surface at each part of the Antarctic. The synergistic fusion of various elevation data and surface structural information enables the output DEMs in the rugged area to retain the fine details of surface morphology. The effective control of the adverse edge effects and the use of the blending weight functions in the data merging minimized the discontinuities between different types of data, resulting in a complete, seamless, and topographically consistent DEM throughout the Antarctic.

CHAPTER 6

CONTINENTAL SCALE DEM AND QUALITY ASSESSMENT

6.1 Seamless continental scale DEM

Having the newly developed and improved algorithms and techniques along with available GIS functions, we successfully constructed a seamless and high-resolution DEM over the Antarctic by integrating a variety of topographical data in a GIS environment (Liu et al., 1999). The final products include three complete DEM grids, respectively with a grid spacing of 200 m, 400 m and 1 km. Each DEM grid covers the entire Antarctic continent and its surrounding offshore ocean area in a polar stereographic projection, and height values are provided with reference to both OSU91A geoid and WGS84 ellipsoid.

Fig. 6.1 is a hill-shaded image constructed from our DEM with 1 km grid spacing at the continental scale. This image clearly shows the overall topographical configuration of the Antarctic, including the ice divides, glacial drainage basins, ice shelves, ice rises, coastal margins, mountain chains and volcanoes. It also captures more subtle features, such as the McMurdo Dry Valley, Lake Vostok, ice flow lines on the Ross Ice Shelf and Ronne-Filchner Ice Shelf, and textured structures probably related to subglacial topography. This image also demonstrates that the data integration and merging methods used in this research work very well. Zoom-in inspection of this image found no data

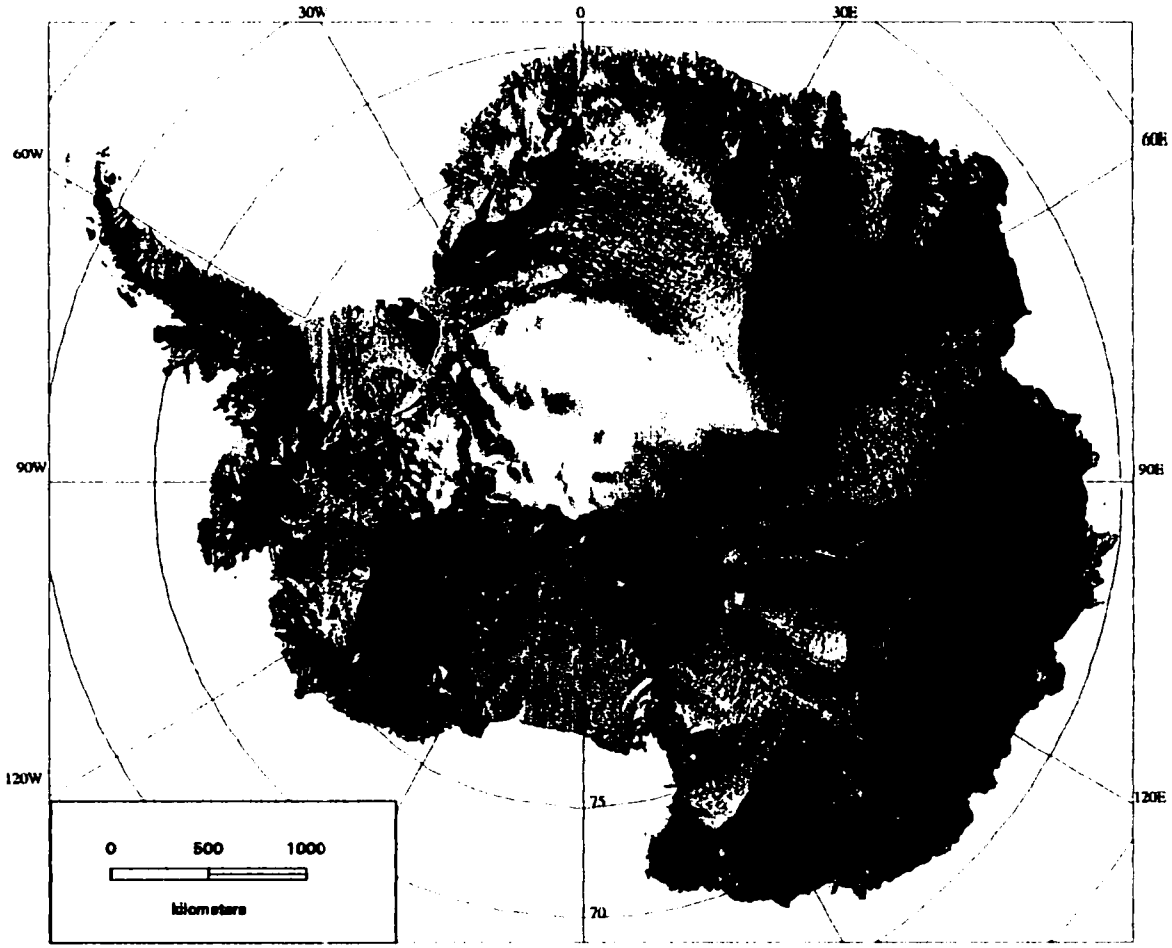


Fig. 6.1: Shaded relief image of the DEM at continental scale

"cliffs" or unnatural, abrupt artifacts.

Also, ice drainage basins and ice flow lines were extracted from the DEM as a simple validation of the topological consistency. First, the high frequency topographical components of the DEM were removed by a low-pass Gaussian filter with a window of a 25 km radius. The ice divides are delineated based on the ridgelines between ice basins. Ice flows along the direction of maximum surface slope under the influence of gravity, and the flow direction can be represented by the aspect of the terrain surface, namely the direction of the steepest descent in surface elevation. Based on this principle, an algorithm is developed to derive ice flow lines. A series of seed points are first generated in both sides of the ice divides within a narrow buffer zone, and then the ice flow line is traced downslope from each seed point in a series of small steps of walk distance Δd (3 km), in the direction of the surface aspect. At each step, the flow line follows the true angle of the aspect at the start point, and the location of the end point for each walk step is determined by the aspect and walk distance. To make the algorithm robust to random errors in aspect calculation, the walk direction of the previous step is stored in memory. If the walk direction at the current step is different from the previous step by 45°, the previous walk direction is used. The tracking procedure for each seed point is terminated when the flow line reaches the coastlines or meets other flow lines (Fig. 6.2). The ice flow lines created from this algorithm is much better than that derived by the ARC/INFO standard flow line extraction algorithm (Jenson and Domingue, 1988; ESRI, 1991b).

The ice divides and ice flow lines derived from our DEM compare favorably with the ice drainage map in Drewry (1983). Visual inspection shows that the locations and directions of the flow lines are consistent with the terrain features in satellite images, and

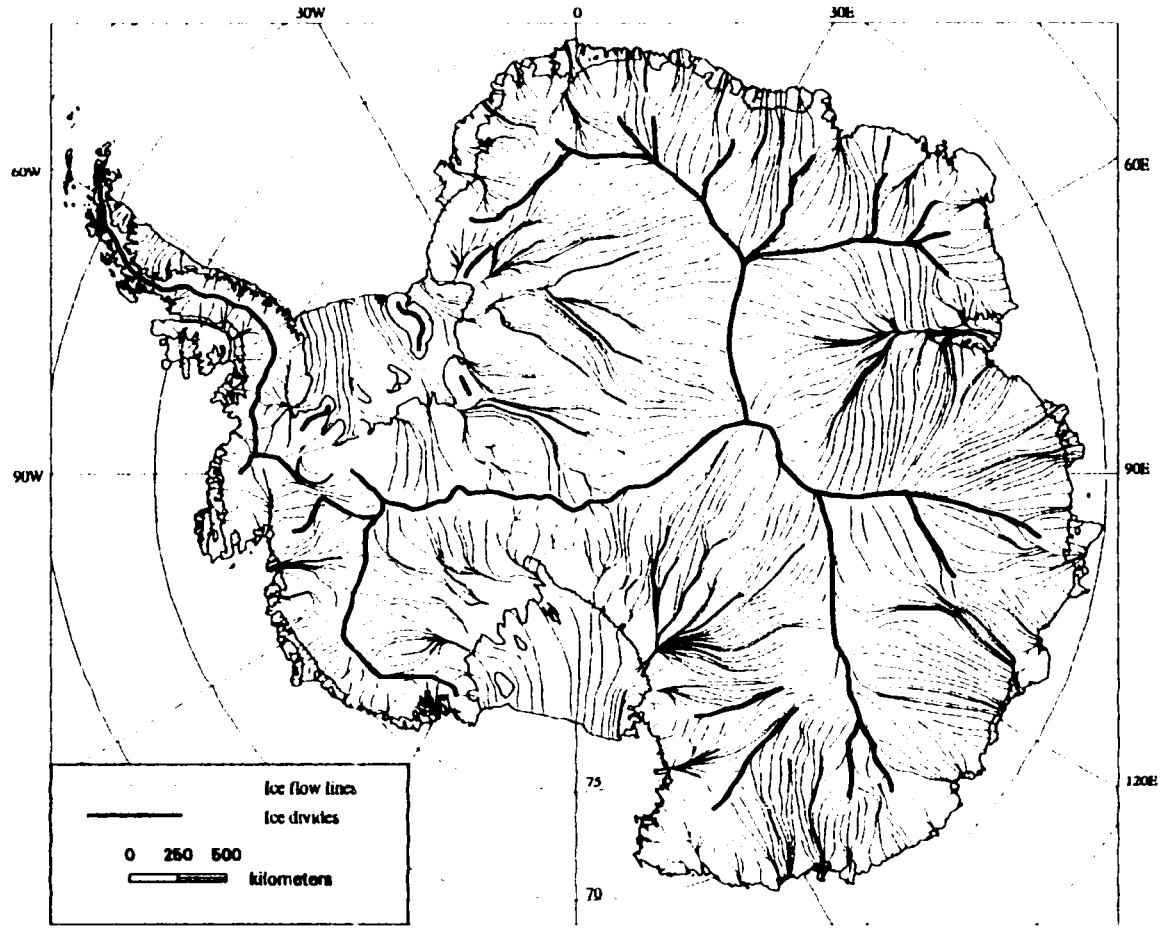


Fig. 6.2: Ice drainage pattern and ice flow lines derived from Antarctic DEM

most of the flow lines converge into major ice streams or outlet glacier channels as indicated by satellite images. This demonstrates that our DEM product is topologically and morphologically consistent.

In addition to the broad topographical configuration and the true ice drainage pattern, our high-resolution DEM also contains considerable information about small-scale terrain features. Fig. 6.3 and Fig. 6.4 compare the simulated images from the DEM with a Landsat TM image of the Oscar II coast of the Antarctic Peninsula and a DISP satellite photograph (McDonald, 1995) near the Beardmore glacier in the Transantarctic Mountains. The DEM simulated relief images present a clear picture of morphologic features and the surface undulations. The shapes and orientations of major glacial valleys and mountain ranges derived from the DEM (200 m resolution) match very well with what are observed in the high resolution Landsat TM image (30 m resolution) and the DISP satellite photograph (about 150 m resolution). Fig. 6.5 presents the comparison of the DEM simulated image with the Radarsat SAR Quicklook image (100 m resolution) of the Victoria Land in the Transantarctic Mountains. Again, a clear picture of the morphologic features is shown in the DEM derived image. The geometric properties of major glacial valleys and mountain ranges derived from the DEM are in good agreement with the SAR image. The high correlation and close matching between the geometric shapes and orientations of topographic features in the simulated image and those in the real satellite images show the level of topographical detail, as well as the topological consistency of the DEM with the real terrain. Shown in Fig. 6.6 are the perspective views of elevation models of the Anverse Island and Brabent Island of Antarctic Peninsula. Fig 6.7 is a terrain-corrected Radarsat SAR image (25 m resolution) draped on the top of the

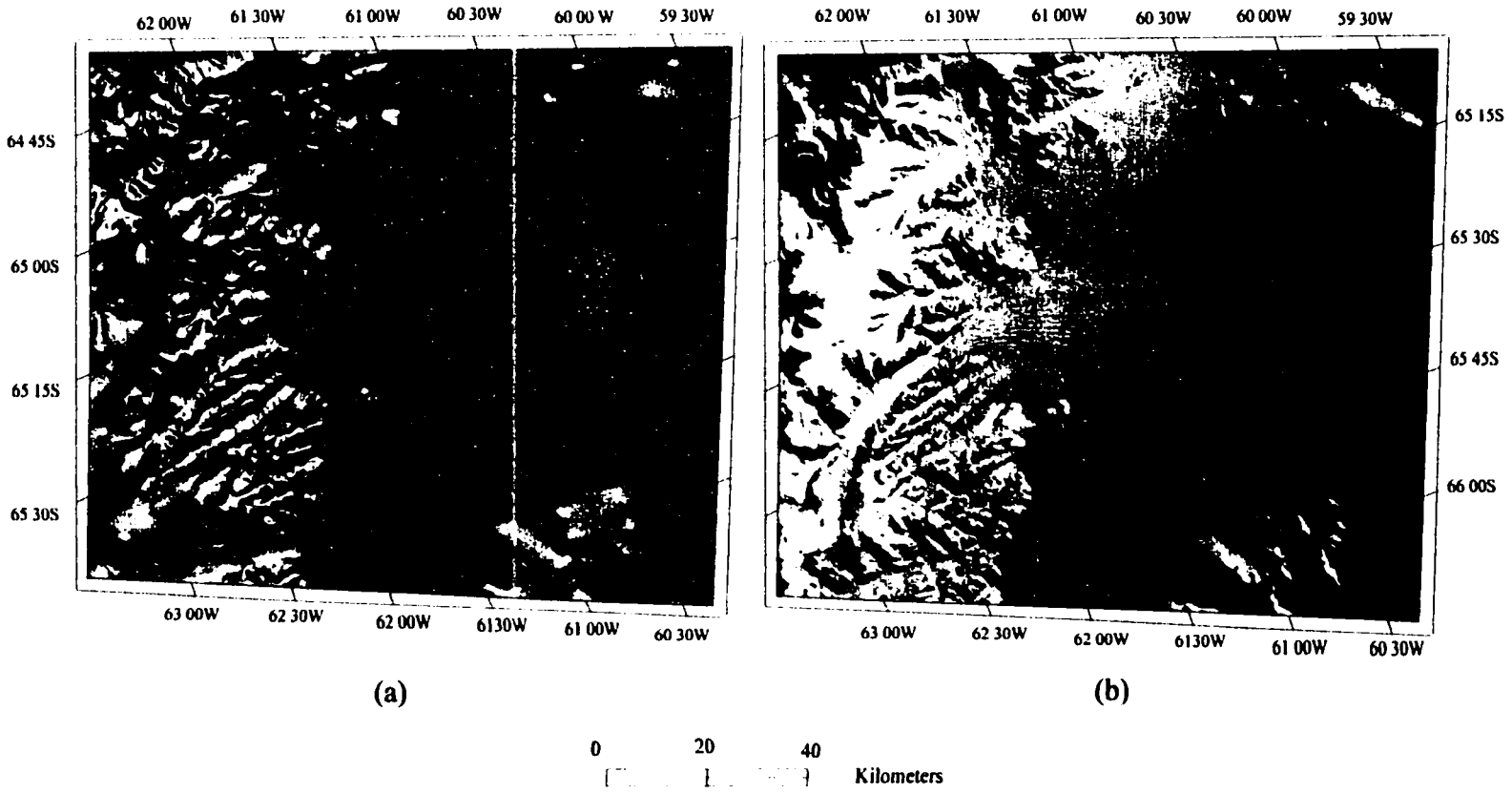


Fig. 6.3: Comparison of the DEM shaded relief image with Landsat TM image in OSCAR II Coast of Antarctic Peninsula. (a) hill-shaded image from DEM (200 m resolution); (b) Landsat TM image (30 m

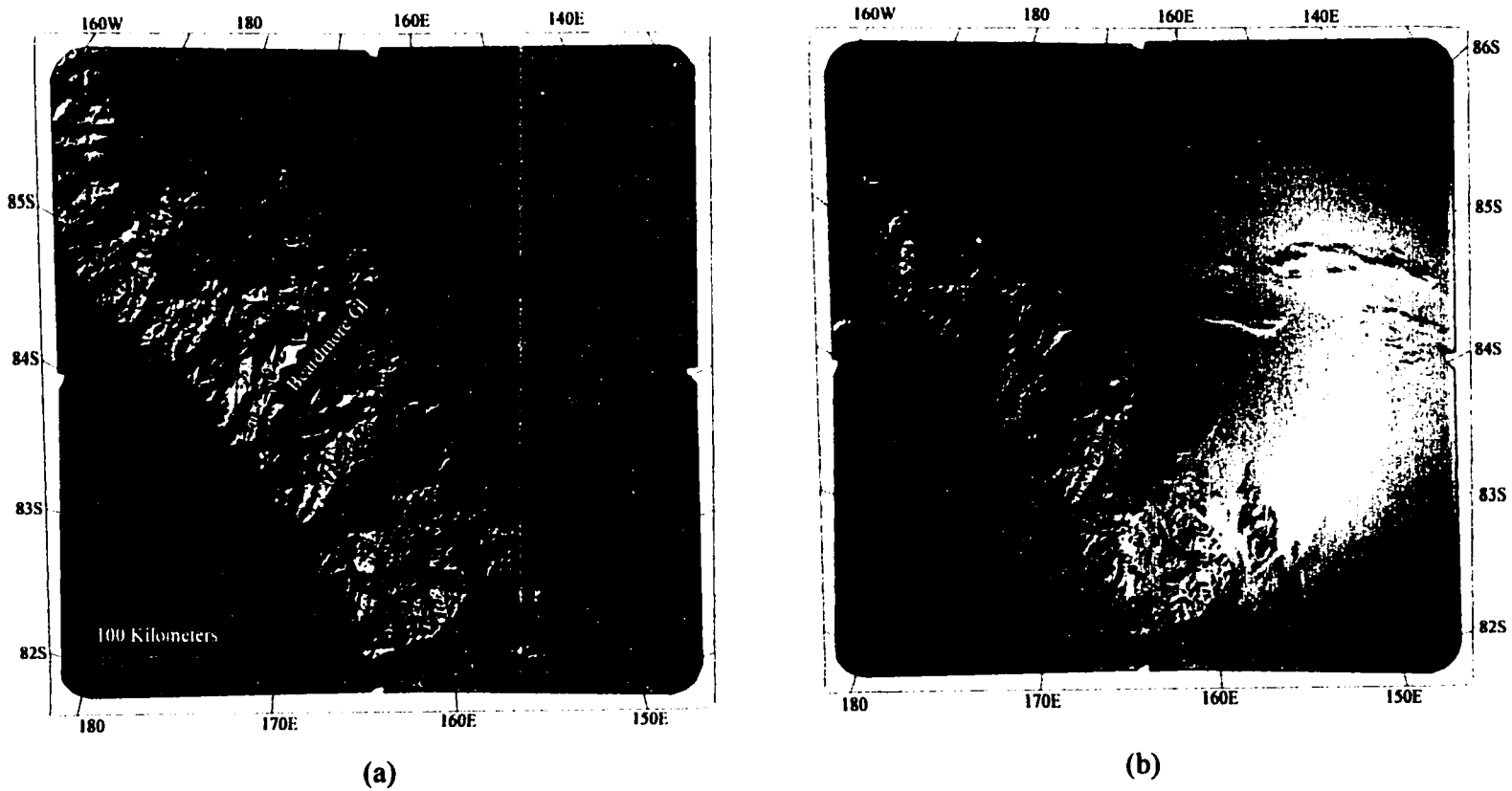
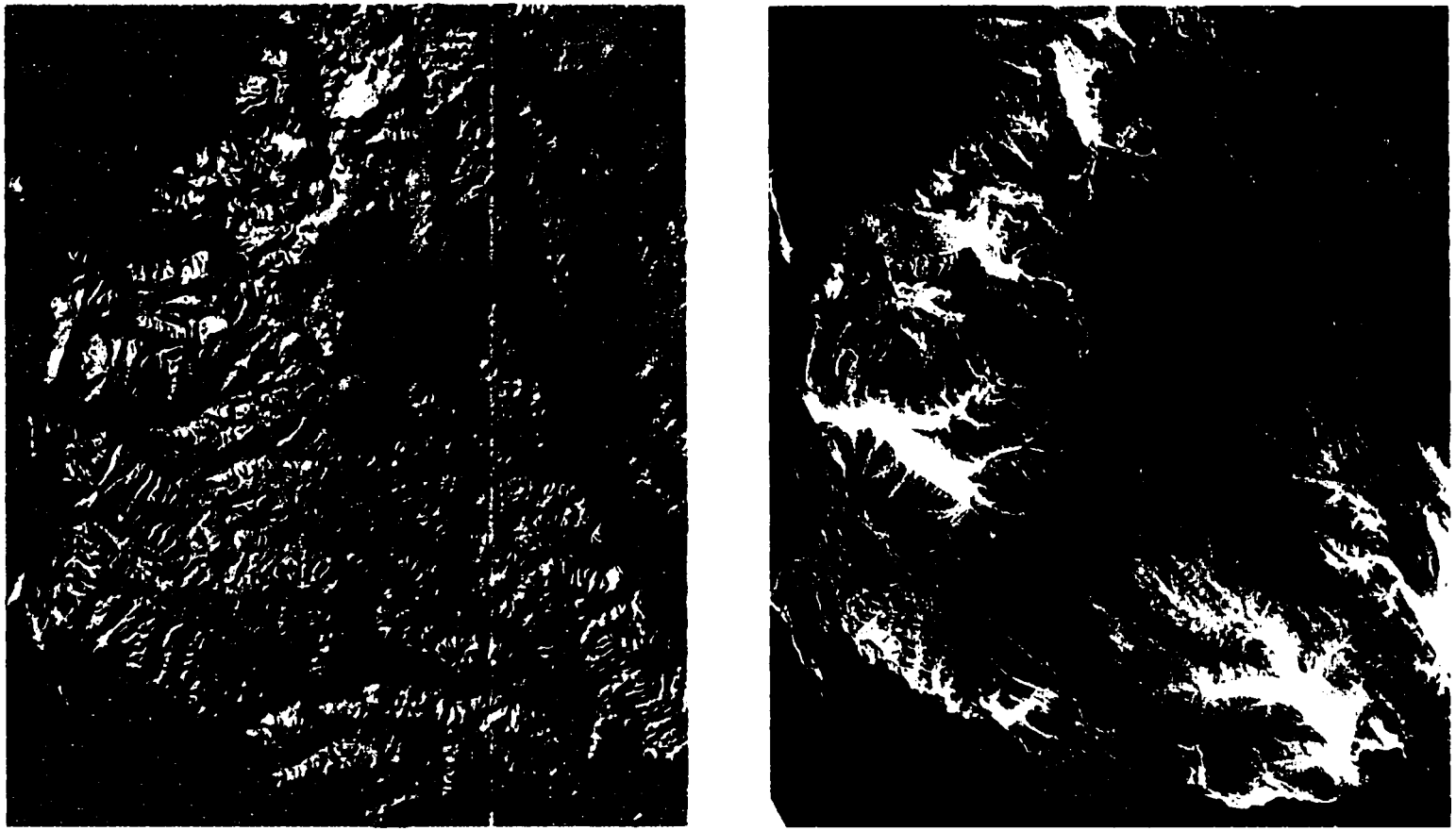


Fig. 6.4: Comparison of the DEM shaded relief image with DISP satellite photograph in Transantarctic Mountains. (a) Hill-shaded image from DEM (200 m resolution); (b) DISP satellite photograph (150 m resolution).



(a) (b)

Fig. 6.5: Comparison of the DEM shaded relief image with Radarsat SAR quicklook image mosaic in Victoria Land of Antarctica. (a) hill-shaded image from DEM (200 m resolution). (b) Radarsat SAR image (100 m resolution)

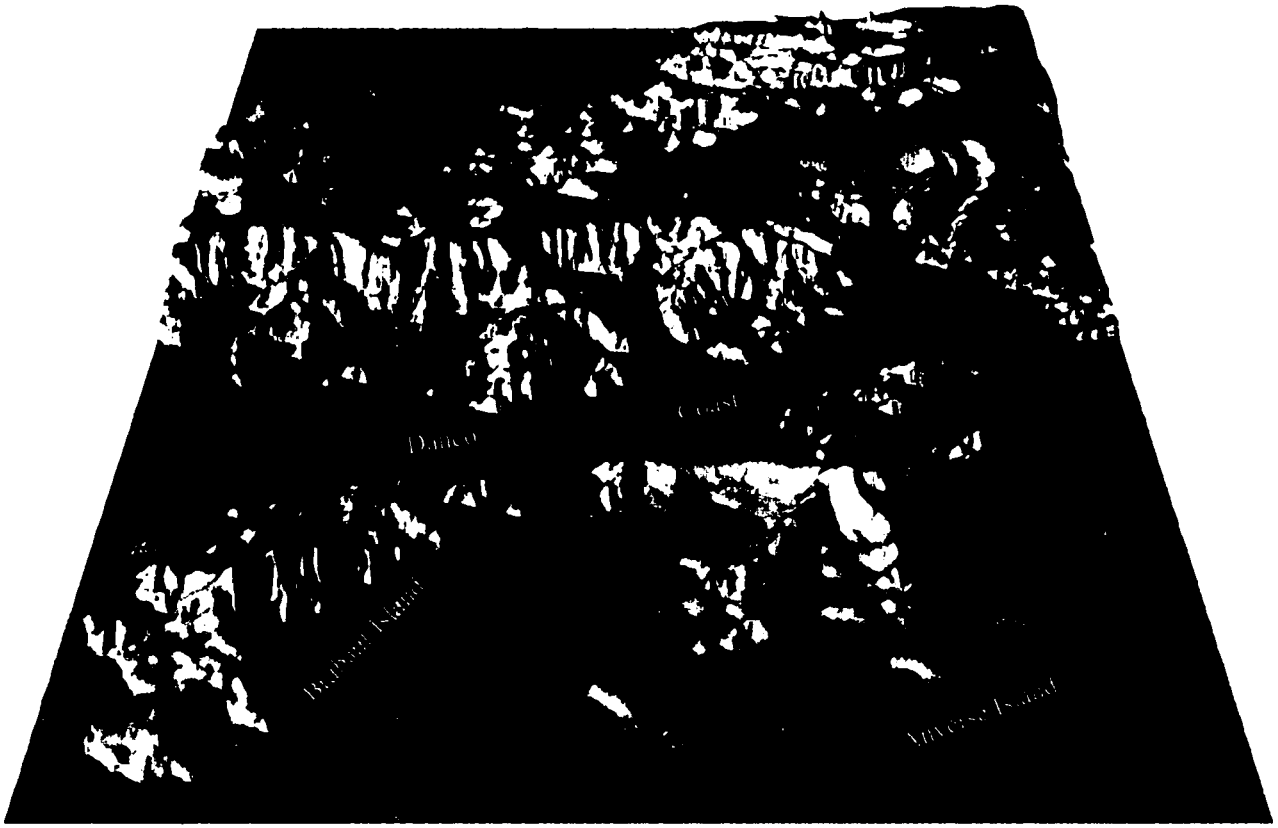


Fig. 6.6: Perspective view of a shaded relief image draped on DEM around Anverse Island and Brabant Island of Antarctic Peninsula

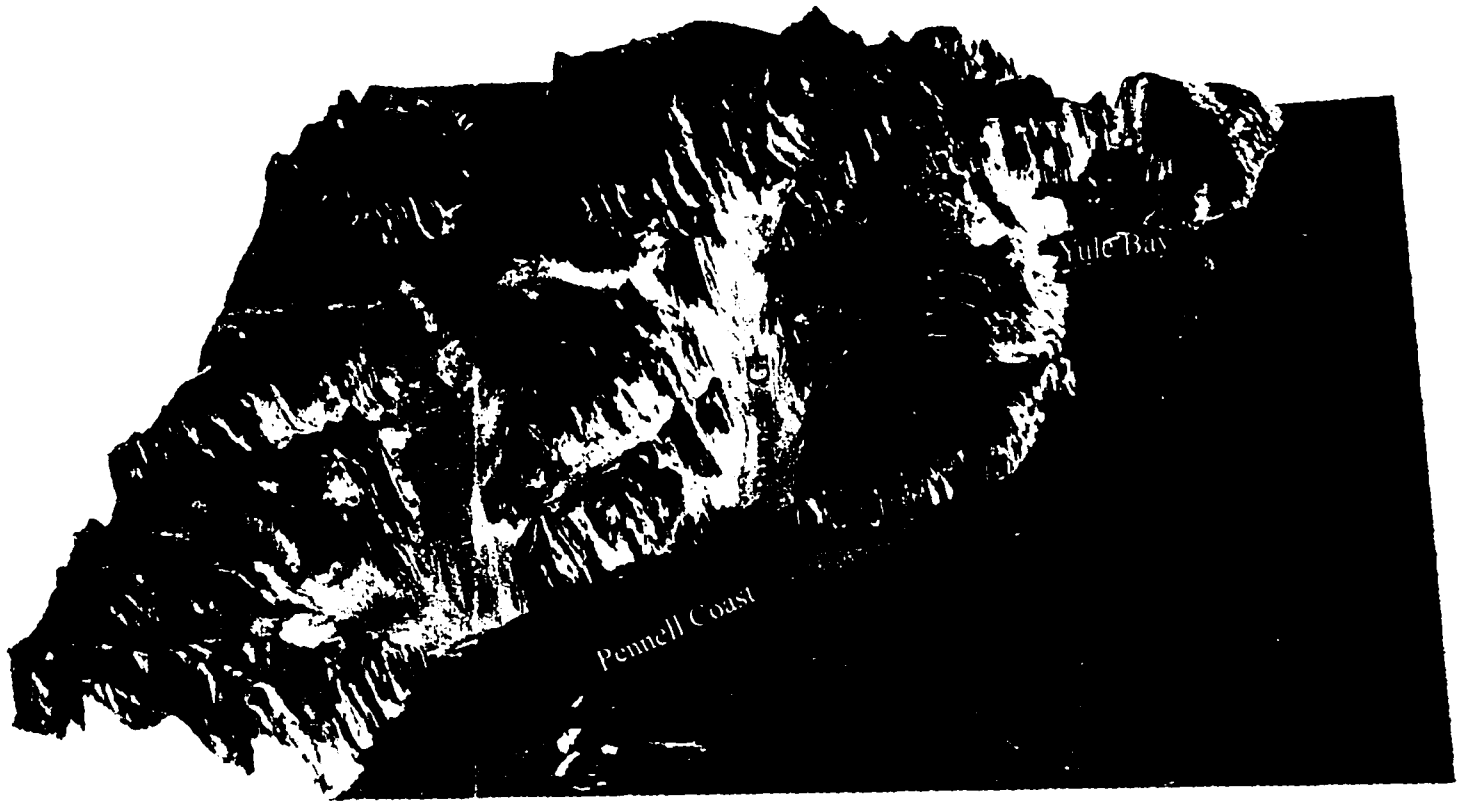


Fig. 6.7: Perspective view of Radarsat SAR image draped on the DEM in Ponnell Coast, Antarctica

DEM of the Pennell Coast. The elevation data of the DEM agree with the terrain features in the SAR image, demonstrating the capability of our DEM in rectifying and visualizing the high-resolution satellite imagery in a rugged area.

6.2 Quality assessment

6.2.1 Comprehensive view of quality

The quality of a DEM is more than just vertical accuracy. The spatial resolution, reliability, and topological consistency are also important aspects. Vertical accuracy measures the correspondence between elevation values of the DEM and the true elevation values of the ground surface. Error is defined, on the other hand, as the deviation of elevation values of the DEM from the true elevation values. Errors can be divided into three types (Thapa and Bossler, 1992): outliers, systematic errors, and random errors. Outliers indicate gross errors or blunders. Systematic errors give consistently higher or lower values than the true values, often referred to as the bias. The reliability is defined as the degree in which the elevation values of DEM are free from blunders and systematic errors. Random errors degrade measurement accuracy by reducing the precision of the measurements. The precision refers to the smallest vertical difference that the measuring instruments can faithfully detect. The spatial (horizontal) resolution indicates how much topographical details that a DEM contains as a whole, often represented by the smallest topographical features that can be extracted from the DEM. The topological consistency is a measure of the correctness of the spatial relationships between elevation values at neighboring nodes of the DEM grid, and it determines whether the terrain surface structural elements and drainage network are correctly represented by the DEM.

In the preceding section, we have demonstrated the topographical information content and topological consistency of our DEM. In the following sections, we will examine the horizontal resolution, positional accuracy, reliability, and vertical accuracy of our Antarctic DEM.

6.2.2 Horizontal resolution and positional accuracy

Though we produced three sets of continental scale DEMs with grid spacings of 200 m, 400 m and 1 km, the real horizontal resolution of each DEM varies from place to place according to the density and scale of the original source data. By diagnosing the information content of original source data, the horizontal resolution of our DEM is estimated at about 200 m in the Transantarctic Mountains and Antarctic Peninsula, and about 400 m in the sloped coastal regions. For the ice shelves and the inland ice sheet covered by satellite radar altimeter data, the horizontal resolution remains about 5 km. Where the airborne radar sounding data were used, the horizontal resolution is estimated at 1 km. The flat plateau inside 81.5°S is covered by a small-scale contour coverage, and the horizontal resolution is estimated at about 10 km.

Positional accuracy of the DEM is also evaluated in a number of selected areas by comparing the simulated images with orthorectified high-resolution Landsat image and slant range Radarsat SAR images. In general, the absolute planimetric accuracy of topographic features extracted from the DEM is better than 100~300 m. An exception is the Ellsworth Mountains. Comparison of the simulated relief image from the DEM with an orthorectified Landsat image and Radarsat SAR images in this region revealed that the source contours of the ADD have 3~5 km systematic positional offsets (Fig. 3.7),

although the shapes of terrain derived from the DEM closely match the satellite images. Through close comparison of the digitized cartographic data with original paper topographical maps, we verified that the digitizing errors are negligible for the cartographic data in the ADD. The positional errors were most likely caused by the poor ground control and inaccurate navigation technique during the 1960s when the topographic maps were produced. The positional errors were corrected by a warping technique based on several selected tie points.

6.2.3 Reliability and vertical accuracy

Three complementary approaches have been used to increase the reliability of our DEM product, including blunder detection in source data, interpolation process verification, and DEM product validation. Owing to extensive and rigorous error checking and correction operations on input source data and the carefully designed interpolation methods, our DEM is free of gross errors or fictitious artifacts. The cross-validation between the redundant topographical data sets greatly reduced the possibility of systematic errors. Therefore, we have a very high confidence in the reliability of our final DEM product.

Absolute vertical accuracy varies with location according to the source data quality and the terrain complexity. The accuracy assessment was performed by comparing the DEM with a very large-scale topographic map, GPS, geodetic leveling, and GPS-navigated airborne radar data. In the accuracy analysis, the mean (m), standard deviation (std) and Root Mean Squared Error (RMSE) are used to quantify the nature and magnitude of errors:

$$d_i = h_i - h_i^{ref} \tag{6.1}$$

$$m = \frac{\sum d_i}{N} \tag{6.2}$$

$$std = \sqrt{\frac{\sum (d_i - m)^2}{N}} \tag{6.3}$$

$$RMSE = \sqrt{\frac{\sum (d_i)^2}{N}} \tag{6.4}$$

where d_i is the height difference; h_i is the elevation values of the DEM; h_i^{ref} is the elevation value of the reference data (ground truth); and N is the number of data points compared. The mean represents a systematic shift between the measured surface and true surface, and the standard deviation is the dispersion of measurement errors. If there is no systematic error, namely, the mean is zero, then the standard deviation equals the RMSE.

To calculate the above accuracy indices, the reference data (ground truth) are needed. Practically, all available high-quality topographical information was included in our DEM generation. However, we achieved a realistic estimate by withholding a small percentage of highly accurate measurements from the interpolation process and use the withheld data as ground truth to obtain the vertical accuracy evaluation. At the final stage, the withheld data are put back into the interpolation process.

For areas where the DEM grid is derived from the ADD cartographic data, the vertical accuracy is mainly influenced by the quality of the original measurements, the map scale, and contour interval. The Transantarctic Mountains, Ellsworth Mountains, the coastal mountains of the Marie Byrd Land and the Queen Maud Land, are primarily covered by 1:250,000 scale contour data with the contour interval of 200 m. The

Antarctic Peninsula is covered by 1:250,000 scale contour data with a contour interval of 250 m. Most coastal regions in East Antarctica are covered by 1:1,000,000 scale contour data with the contour interval of 100 m. The interior of the ice sheet is covered by 1:3,000,000 scale contour data with a contour interval of 200 m. For the Dry Valley region, one of the most rugged areas in the Antarctic, a DEM data set derived from the ADD 1:250,000 scale contour data was compared with the USGS 1:50,000 scale contour data that are assumed accurate to represent the ground truth. The mean of the error is 75 m with a standard deviation of 109 m, and the Root Mean Squared Error (RMSE) is 132 m (Liu and Jezek, 1999).

The elevation values of the DEM were compared with GPS traverses over the Lambert Glacier Basin, Amery Ice Shelf, and the Siple Dome, with GPS navigated airborne Radio Echo Sounding (RES) traverse over the Evans Ice Stream, and with geodetic leveling traverses over the Victoria Land and the Marie Byrd Land (Fig. 6.8 and Table 6.1). The Lambert Glacier Basin GPS traverse route is about 2000 km long, consisting of 73 points with about 30 km intervals (Kiernan, 1998). The elevation values in the DEM derived from ERS-1 radar altimeter data are in good agreement with GPS measurements (Fig. 6.8). The RMSE is only 12 m. The Amery Ice Shelf GPS traverse is selected from a denser kinematic GPS survey data set. The original survey consists of a series of 24 10km x 10km square grids with the accuracy of 15 cm (Phillips et al., 1996). 13 GPS corner points in the middle of the Amery Ice Shelf with a spacing of 10 km were selected for comparison. As shown in Fig. 6.9, the ERS-1 radar altimeter derived elevation values in the DEM agree with the kinematic GPS measurements within 2 m. The Siple Dome GPS traverse is about 120 km long, where the DEM data is derived from the cartographic

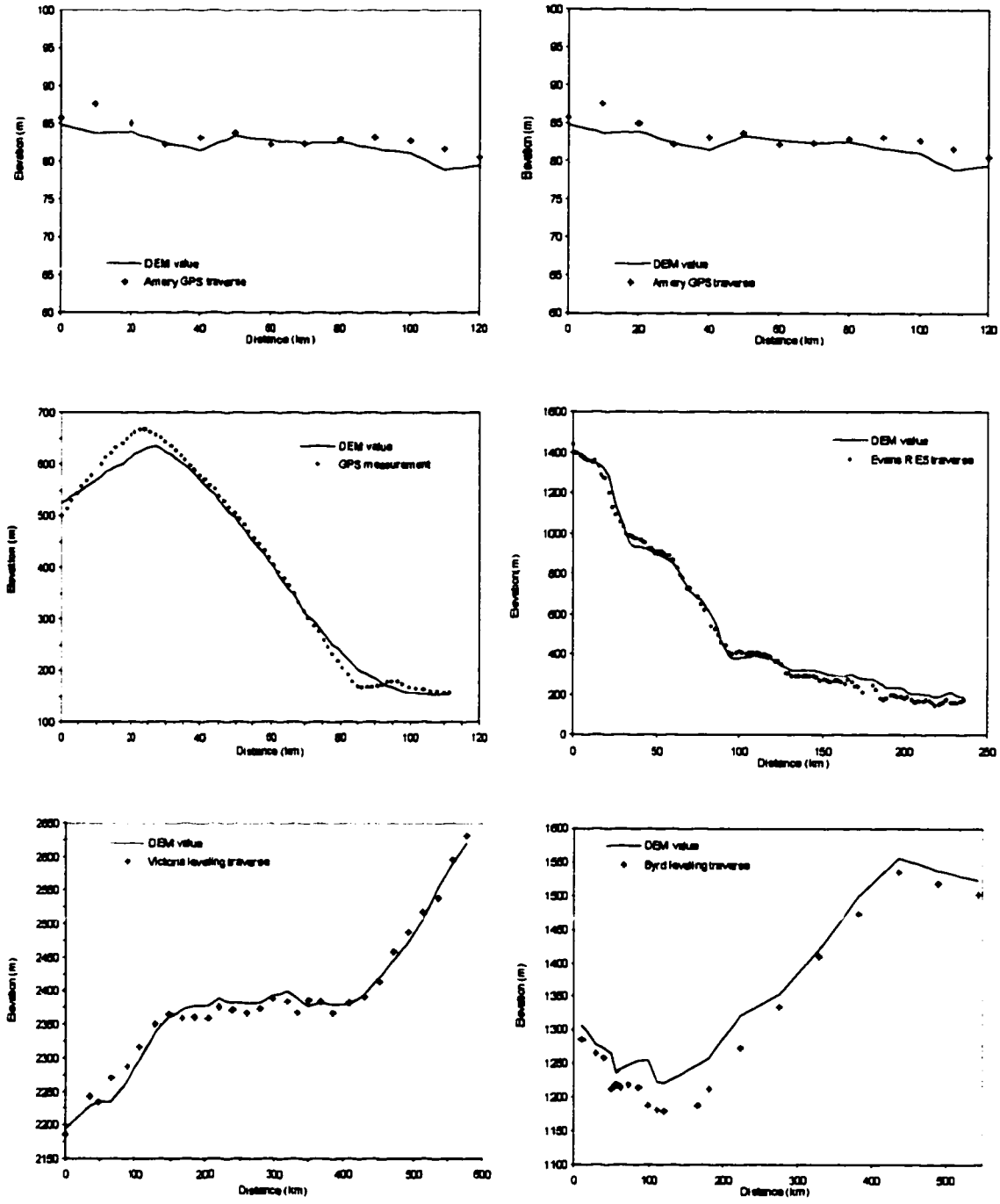


Fig. 6.8: Accuracy assessment of the DEM by comparing with GPS, RES and leveling data. (a) Lambert Glacier Basin GPS traverse; (b) Amery GPS traverse; (c) Siple Dome GPS traverse; (d) Evans Ice Stream RES traverse; (e) Victoria Land geodetic leveling traverse; (f) Marie Byrd Land geodetic leveling traverse.

Traverse	Length km	Mean Error meter	Std. Dev. meter	RMSE meter
Lambert GPS	2000	1	12	12
Amery GPS	120	-1	1	2
Siple Dome GPS	120	-6	18	19
Evans RES	240	15	29	33
Victoria Leveling	600	0	14	14
Byrd Leveling	580	32	16	36

Table 6.1: Vertical accuracy estimates of the Antarctic DEM

data in the ADD. The RMSE of the error along the Siple Dome traverse is about 20 m (Fig. 6.8). The selected Evans Ice Stream airborne RES traverse is about 240 km long. If the airborne radar echo sounding data had been ignored, and only the satellite radar altimeter data were used in this area, the corresponding RMSE is 33 m (Fig. 6.8). The Victoria Land geodetic leveling traverse and the Marie Byrd Land geodetic leveling traverse are extracted from the ADD spot height coverage. The oversnow geodetic leveling data are reported accurate to better than 10 m (Drewry, 1983). The comparisons of the DEM data derived from ERS-1 radar altimetery with these two leveling traverses resulted in the RMSE values of 14 m and 36 m (Fig. 6.8). There is no systematic error between the ERS-1 radar altimeter derived elevation data and the Lambert GPS traverse, Amery Ice Shelf GPS traverse, and Victoria Land geodetic leveling traverse, indicated by very small mean error values (Table 6.1). Along the Evans RES traverse and the Marie Byrd Land geodetic leveling traverse, the overall overestimate of surface height is

indicated by the positive mean errors of 17 m and 32 m.

In summary, the accuracy of the DEM is estimated about 100~130 m over the rugged mountainous areas, about 20~50 m in the highly sloped coastal regions, better than 2m in the ice shelves, and better than 15 m in the low-slope ice sheet and about 35 m in the relatively rough and steep ice sheet (0.65~1.0°).

It should be noted that the above accuracy assessments are only based on selected locations where some type of independent data with a higher order of accuracy exist. However, there remain some regions of the Antarctic where no reliable topographic data exist. Correspondingly, the accuracy of the DEM in these areas is suspect. For example, the best existing topographic data for the flat plateau inside the 81.4°S circle is a 1:3,000,000 scale map. The Prince Charles Mountains in the Lambert glacier basin are only covered by 1:1,000,000 scale map, and parts of Pensacola Mountains and Shakleton Range on the coast of Ronne-Filchner Ice Shelf are covered by 1:1,000,000 and 1:3,000,000 scale map.

6.3 Error pattern analysis

DEMs can be used to derive a variety of new variables and parameters. Previous research shows that the accuracy of derived variables is affected, not merely by the magnitude of DEM errors and the algorithms applied to derive these variables, but also by the spatial structure of DEM errors (Lee et al. 1992; Hunter and Goodchild 1997; Henvelink 1998). A proper understanding of the spatial pattern of DEM error is a basic premise in modeling error propagation. As recognized by Hunter and Goodchild (1997)

and Monckton (1994), the lack of knowledge about the spatial structure of DEM errors has hampered studies that track and model error propagation.

To explore the error pattern of our DEM product, an empirical analysis of a test site was conducted (Liu and Jezek, 1999). The test site is a portion of the McMurdo Dry Valley region in the Transantarctic Mountains, and the terrain is characterized by rugged glacial landforms (Fig. 6.9). The DEM data set under investigation is derived from a 1:250,000 topographical map with a contour interval of 200 m. The reference data are extracted from USGS topographical maps at 1:50,000 scale, with a contour interval of 50 m. The map was photogrammetrically generated in 1975 from U.S. Navy aerial photographs taken in 1970. It qualifies as a benchmark for the DEM accuracy assessment because it originates from an independent data source and has a much larger scale and higher accuracy in comparison with the data source used to create the DEM data set (Fig. 6.9).

For each individual data point i, we define arithmetic error and absolute error as follows:

Arithmetic error:

$$e_i = H_i^{DEM} - H_i^{REF} \tag{6.5}$$

Absolute error:

$$|e_i| = |H_i^{DEM} - H_i^{REF}| \tag{6.6}$$

where H_i^{DEM} is the elevation value in the DEM at point i; H_i^{REF} is the elevation value of the reference data at the corresponding point i. Fig. 6.10 shows the spatial distribution of the absolute error term defined in (6.6). It is quite clear that the magnitude and gradient

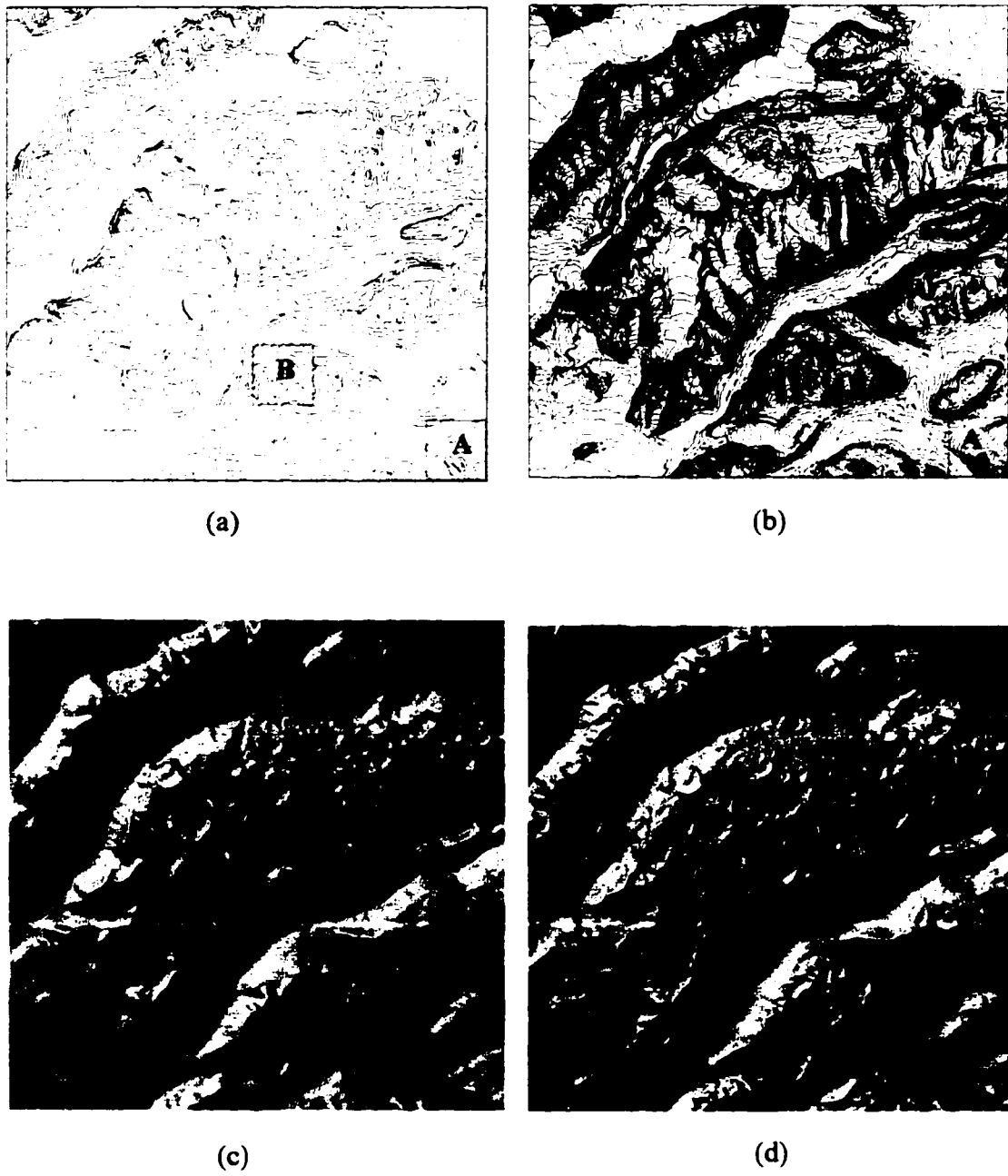


Fig. 6.9: Comparison of McMurdo Dry Valley DEM data with reference data. (a) contours of DEM data source from 1:250,000 USGS topographic map; (b) contours of reference data from 1:50,000 USGS topographic map; (c) analytical hillshading of DEM data; (d) analytical hillshading of reference data.

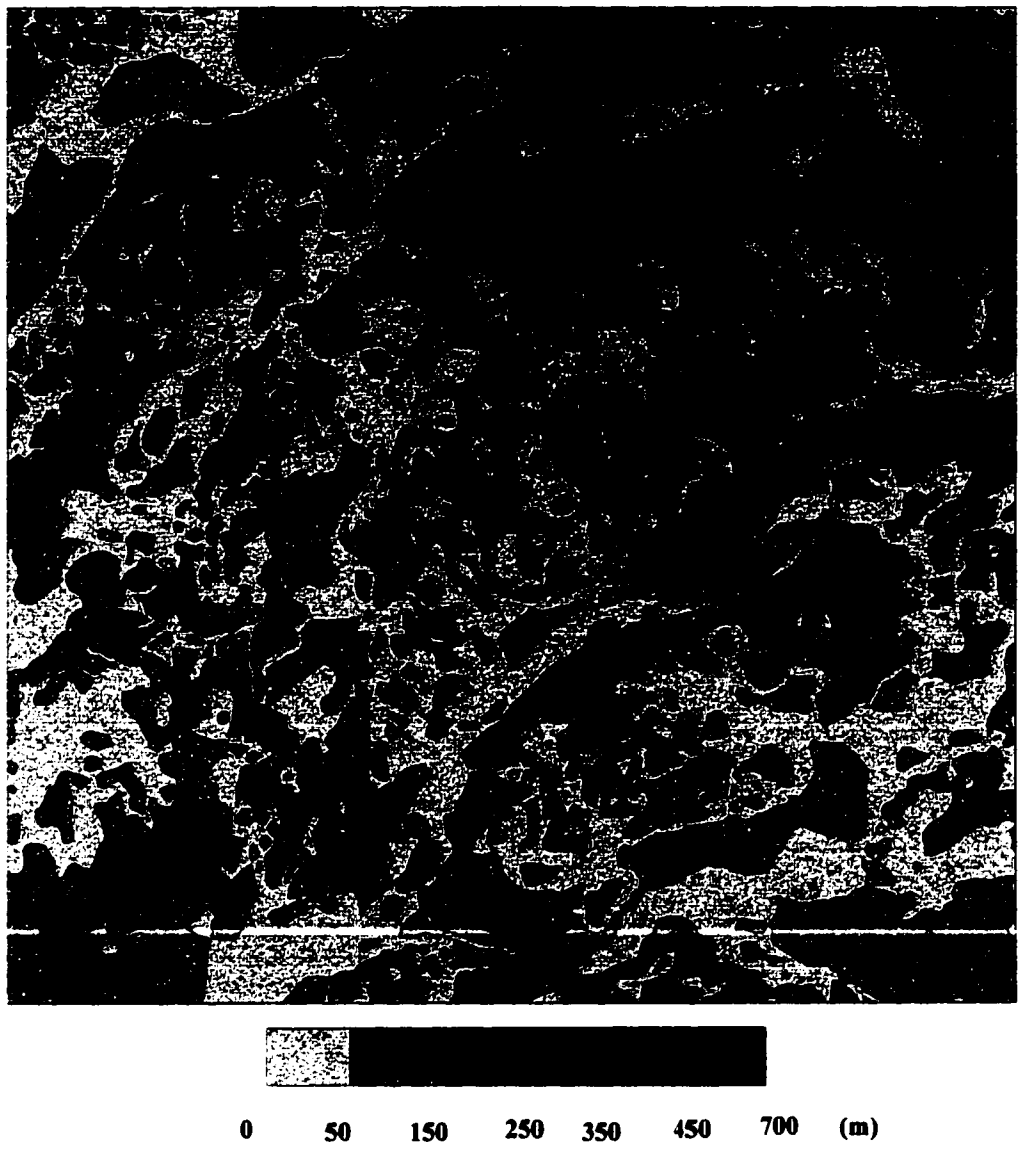


Fig. 6.10: Spatial distribution of absolute errors, with the base contour line of 50 m and contour interval of 100 m

of errors is much smaller in the flat floors of the glacier or ice-free valleys than in the rugged slopes of the mountains.

The directional variograms and Fourier analysis are introduced to quantify and characterize the spatial structure of the DEM errors. The experimental variogram is half the average of the squared difference between pairs of data values of the regionalized variable, in our case, the DEM error term, at a spatial separation lag $\boldsymbol{h}$ (Isaaks and Srivastava 1989; Deutsch and Journel 1992):

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{l=1}^{N(h)} [e(x_l) - e(x_l + h)]^2 \tag{6.7}$$

where $\hat{\gamma}(h)$ is the variogram value at spatial separation lag $\boldsymbol{h}$; $N(h)$ is the number of observation pairs at lag $\boldsymbol{h}$; and $e(x_l)$ and $e(x_l + h)$ are the error values at points x_l and $x_l + h$.

Similarly, the covariogram (covariance function) can be defined as follows (Isaaks and Srivastava 1989; Deutsch and Journel 1992):

$$\hat{C}(h) = \frac{1}{N(h)} \sum_{l=1}^{N(h)} [e(x_l) - \bar{e}][e(x_l + h) - \bar{e}] \tag{6.8}$$

where $\hat{C}(h)$ is the covariogram value at lag $\boldsymbol{h}$; $\bar{e}$ is the mean of all the observed sample values at the lag $\boldsymbol{h}$; and the other parameters are the same as in (3).

Theoretically, the variogram and covariogram are related by

$$\gamma(h) = C(0) - C(h) \tag{6.9}$$

To examine the spatial structure of the DEM errors, both variograms and covariograms are calculated based on the irregular error data set along six directions with a lag distance step of 200 m using the algorithms in GSLIB (Deutsch and Journel 1992).

To ensure sufficient pairs for clear variograms and covariograms, a directional tolerance of 5° and a lag tolerance of 50 m are used in the calculations.

The existence of spatial autocorrelation of the DEM error is clearly demonstrated by the variograms (Fig. 6.11). The most striking characteristic observed from the experimental variograms is the anisotropic autocorrelation pattern of the DEM error at the medium and far lag distances. Both directional variograms and covariograms begin to diverge at the lag distance of about 500 m (Fig. 6.11). The anisotropic pattern of the error autocorrelation is a mixture of the geometric and zonal anisotropy (Isaaks and Srivastava 1989). For instance, variograms in the azimuthal directions of 60° and 90° have similar sill values, but different ranges, indicating "geometric anisotropy", while variograms in azimuthal directions of 150° and 90° have similar ranges but different sills, indicating "zonal anisotropy". The greatest autocorrelation direction is along azimuth 60° with an autocorrelation range of about 5 km. The perpendicular azimuthal direction of 150° approximately defines the minimum autocorrelation direction with an autocorrelation range of about 3 km, resulting in an anisotropy ratio of about 1.7:1.

The two-dimensional Fourier analysis technique is utilized in two ways. First, it is used as an efficient computational method to derive a 2D autocorrelation pattern of the DEM error based on the Wiener-Khintchine autocorrelation theorem (Weaver 1983; Pardo-Iguzquiza and Chica-Olmo 1993). Second, it is used to decompose the spatial variation of the terrain surface into sinusoidal waves of different frequency in a frequency domain based on the Parseval-Rayleigh theorem (Weaver 1983; Pardo-Iguzquiza and Chica-Olmo 1993) and to identify the frequency components responsible for the DEM error pattern.

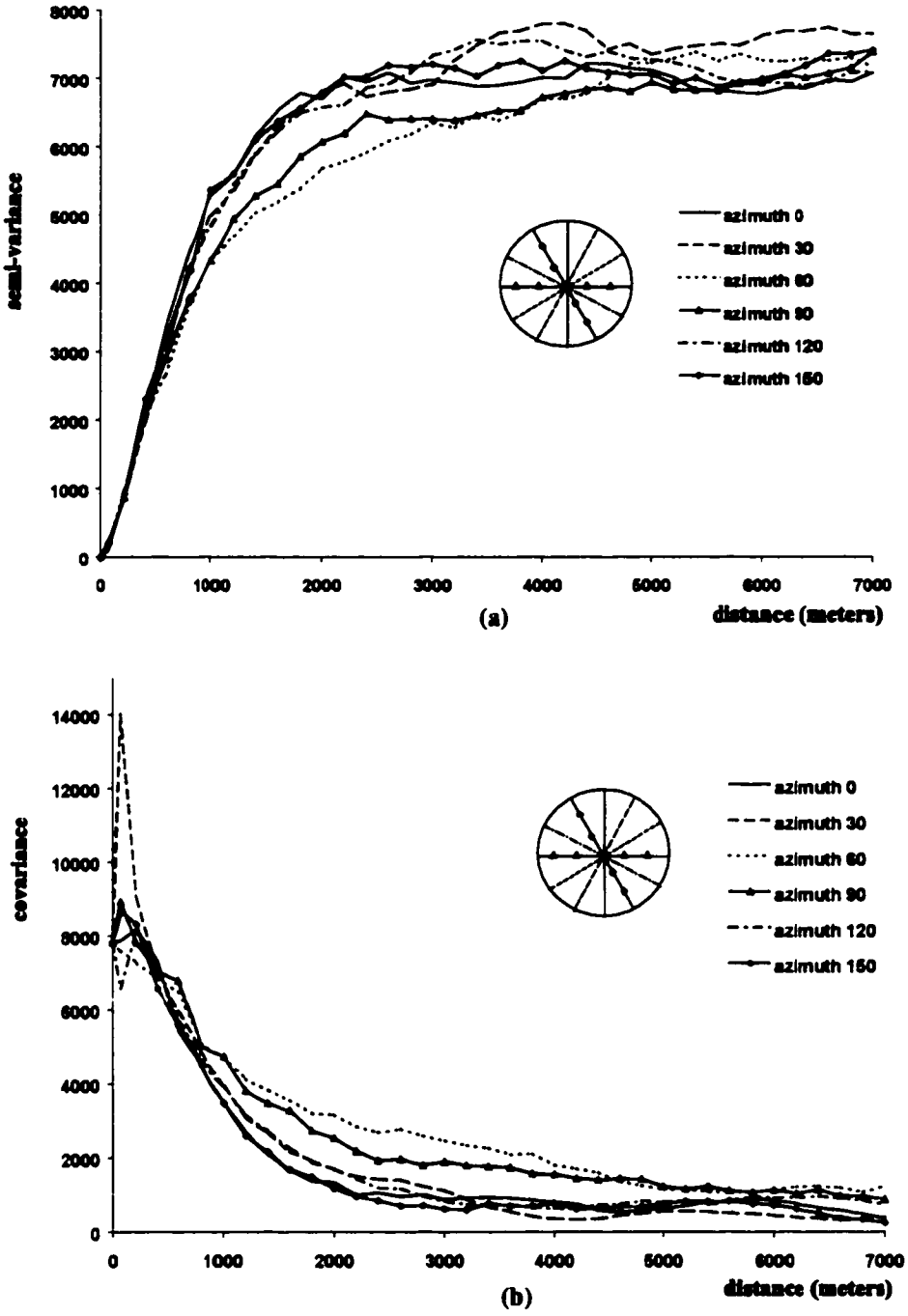


Fig. 6.11: Directional variograms and covariograms of DEM errors. (a) Variograms of DEM errors, calculated using a lag distance step of 200 m. (b) Covariograms of DEM errors, calculated using the same lag distance step as the variograms.

Mathematically, the Fourier transformation of a function f(x,y) is defined as (Gonzalez and Wintz 1987):

$$F\{f(x,y)\} = F(u,v) = \int_{-\infty}^{\infty}\int f(x,y)\exp[-j2\pi(ux+vy)]dxdy \qquad (6.10)$$

the inverse Fourier transformation as:

$$F^{-1}\{F(u,v)\} = f(x,y) = \int_{-\infty}^{\infty}\int F(u,v)\exp[j2\pi(ux+vy)]dudv \qquad (6.11)$$

power spectral density as:

$$P(u,v) = |F(u,v)|^2 = R^2(u,v) + I^2(u,v) \qquad (6.12)$$

autocorrelation function as:

$$C(\xi,\eta) = \int_{-\infty}^{+\infty}\int f(x,y)f(x+\xi,y+\eta)dxdy \qquad (6.13)$$

In equations (6.10)-(6.13), $F\{\}$ is the forward Fourier transform operator; $F^{-1}\{\}$ is the inverse Fourier transform operator; f(x,y) is a two-dimensional spatial variable, representing the DEM error term or the elevation values of the terrain surface in our case; $j=\sqrt{-1}$ imaginary unity; (u,v) is the frequency variable; $F(u,v)$ is the Fourier transformation of f(x, y); $R(u,v)$ and $I(u,v)$ are respectively the real part and imaginary part of the complex $F(u,v)$; $P(u,v)$ is the power spectral density of $F(u,v)$; and $C(\xi,\eta)$ is the autocorrelation function of $f(x,y)$.

Wiener-Khintchine's autocorrelation theorem states that the autocorrelation function and power spectral density function forms a Fourier transformation pair (Weaver 1983; Press et al. 1992), namely:

$$C(\xi,\eta) \Leftrightarrow P(u,v) \tag{6.14}$$

$$C(\xi,\eta) = F^{-1}\{P(u,v)\} = \iint P(u,v)\exp[j2\pi(\xi u + \eta v)]dudv \tag{6.15}$$

Where $\Leftrightarrow$ is the Fourier transform pair.

According to this theorem, the autocorrelation function can be rapidly and efficiently computed by the inverse Fourier transformation of the power spectral density function using the Fast Fourier Transform (FFT) algorithm.

The Parseval-Rayleigh theorem relates the Fourier power spectral density to the total variance computed in the space domain as follows (Weaver 1983; Ansoult 1989):

$$\int_{-\infty}^{\infty}\int |f(x,y)|^2 dxdy = \int_{-\infty}^{\infty}\int P(u,v)dudv = \int_{-\infty}^{\infty}\int |F(u,v)|^2 dudv \tag{6.16}$$

Based on the Parseval-Rayleigh theorem, we can identify frequencies with a large power, hence accounting for a large variance component in the terrain surface and indicating a periodic nature of the terrain surface at that particular frequency. The contribution of each frequency to the overall surface variation can be also computed.

Based on the regular DEM error data set, the 2D spatial autocorrelation pattern of DEM errors is derived using the following steps:

- remove the mean of the error (zero frequency or DC term) to minimize leakage and interference between frequency bands;
- apply the Hanning window to reduce the spectral leakage (the ringing and ripples) caused by the data truncation (Weaver 1983);
- perform the forward Fast Fourier Transformation (FFT);
- calculate the power spectral density; and

- compute the autocorrelation by the inverse Fourier transformation of the power spectral density.

The above computational steps are accomplished by C programming, and most of the routines are taken and modified from Press et al. (1992). The spatial autocorrelation of the DEM error is represented as a contour map in Fig. 6.12. A close look at this map reveals the scale-dependent nature of the spatial autocorrelation pattern of the DEM error. In 2D space, iso-autocorrelation lines are circular around the origin within 500 m of spatial separation. It indicates that autocorrelation within this lag distance is isotropic. For lag distances between 500 m and about 1.5 km, iso-autocorrelation lines become concentric ellipses, indicating an apparent anisotropic pattern. The anisotropy ellipse is oriented with its major axis (maximum continuity) parallel to the azimuthal direction of 0°, and the anisotropy ratio to its perpendicular minimum axis is about 1.3:1. When the lag distance is greater than 1.5 km, the orientation of the anisotropy ellipses begins to change, and the main axis is aligned approximately along the azimuthal direction of 60°. The anisotropy ratio becomes larger, approaching to about 1.8:1. In comparison with directional variograms, the 2D autocorrelation pattern derived from Fourier analysis shows more details at the near lag distances, although the results from directional variograms and Fourier analysis are similar at the medium and far lag distance.

The cross-sectional profiles of the autocorrelation surface along the directions of maximum and minimum continuity are shown in Fig. 6.12. The variation in the density of iso-autocorrelation lines and the slope of profiles suggests that the autocorrelation of DEM errors declines faster within a short lag distance over 1.5 km, and does not continue to drop as quickly within a lag distance from 1.5 km to about 5 km.

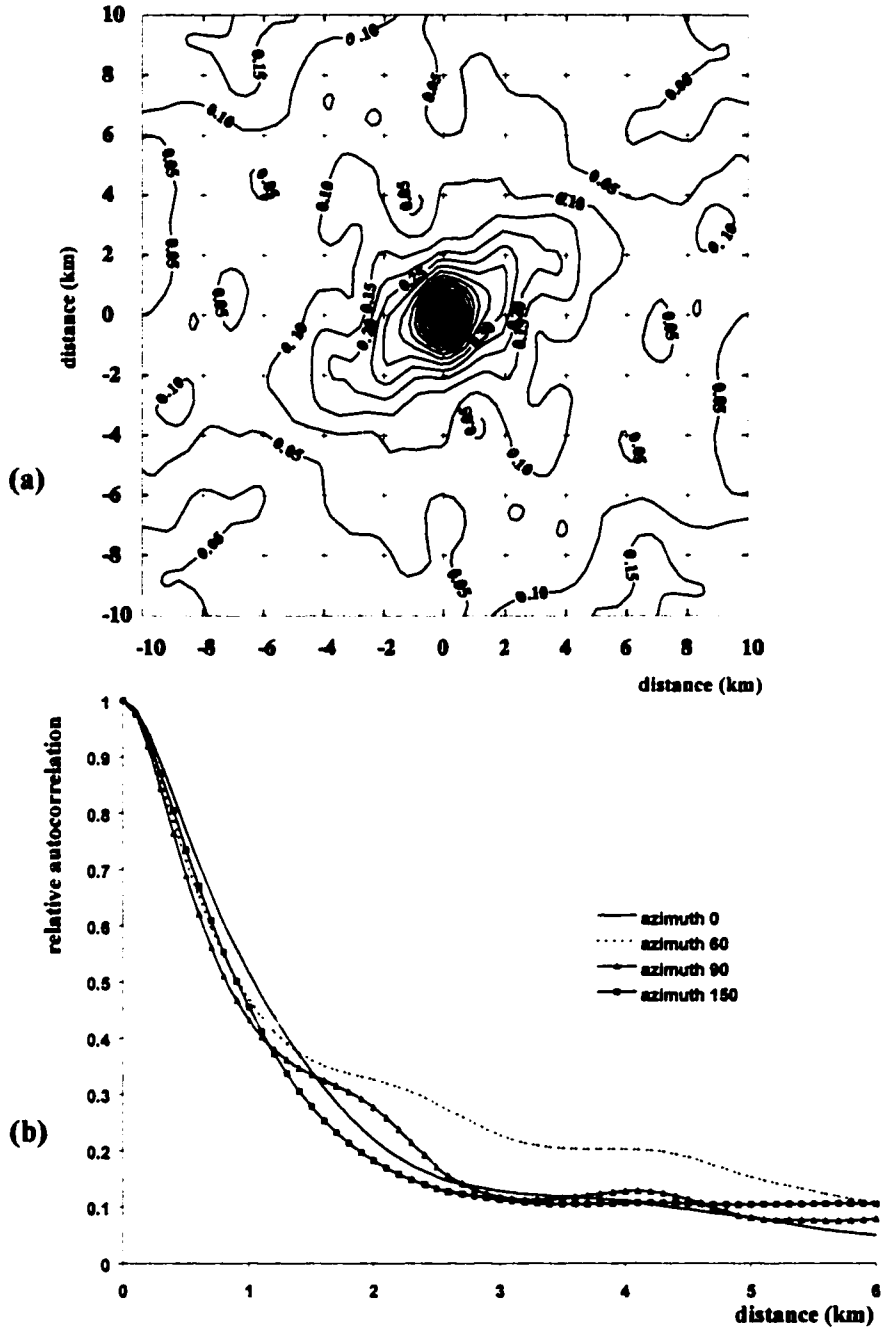


Fig. 6.12: Spatial autocorrelation structure of the DEM errors. (a) 2D anisotropic autocorrelation pattern of DEM errors derived from Fourier analysis; (b) cross-sectional profiles of autocorrelation along maximum and minimum continuity directions.

By visually inspecting Fig. 6.12 in comparison with the terrain surface shown in Fig. 6.9, it becomes clear that the scale-dependent and nested anisotropic autocorrelation pattern of DEM errors is closely related to the morphological features of the terrain topography. At a large scale, the terrain surface under study is dominated by a number of glacier valleys and mountain ranges, which run approximately parallel to the azimuthal direction of 60°. This coincides with the maximum autocorrelation direction of DEM error at a medium and long lag distances. Examining the mountain ranges at a smaller scale, we see that the prevailing landform features are the narrow aretes (eroded ridges) and cirque troughs. The overwhelming majority of these small ridges and troughs are oriented approximately in the azimuthal direction of 0°, which is consistent with the main axis of autocorrelation ellipses of DEM errors at a small lag distance.

The power spectral density is calculated to examine the frequency content of the terrain surface. Two cross-sectional profiles are shown in Fig. 6.13. The displayed spectral powers at different frequencies are in unit of decibel (dB) (Gonzalez and Wintz 1987) calculated by:

$$D(u,v) = \log(1 + P(u,v)) \tag{6.17}$$

where $D(u,v)$ is the displayed power at the frequency (u,v); and $P(u,v)$ is the power spectral density calculated by equation (6.8).

The power spectral profile in the azimuthal direction of 150° shows one strong power peak at a frequency of about 11 cycles/51.2 km. This is the average frequency of the periodic arrangement of glacier valleys and mountain ranges in this direction. Therefore, the average wavelength is about 5 km. Half the wavelength (2.5 km) approximately corresponds to the autocorrelation length of the minimum axis of

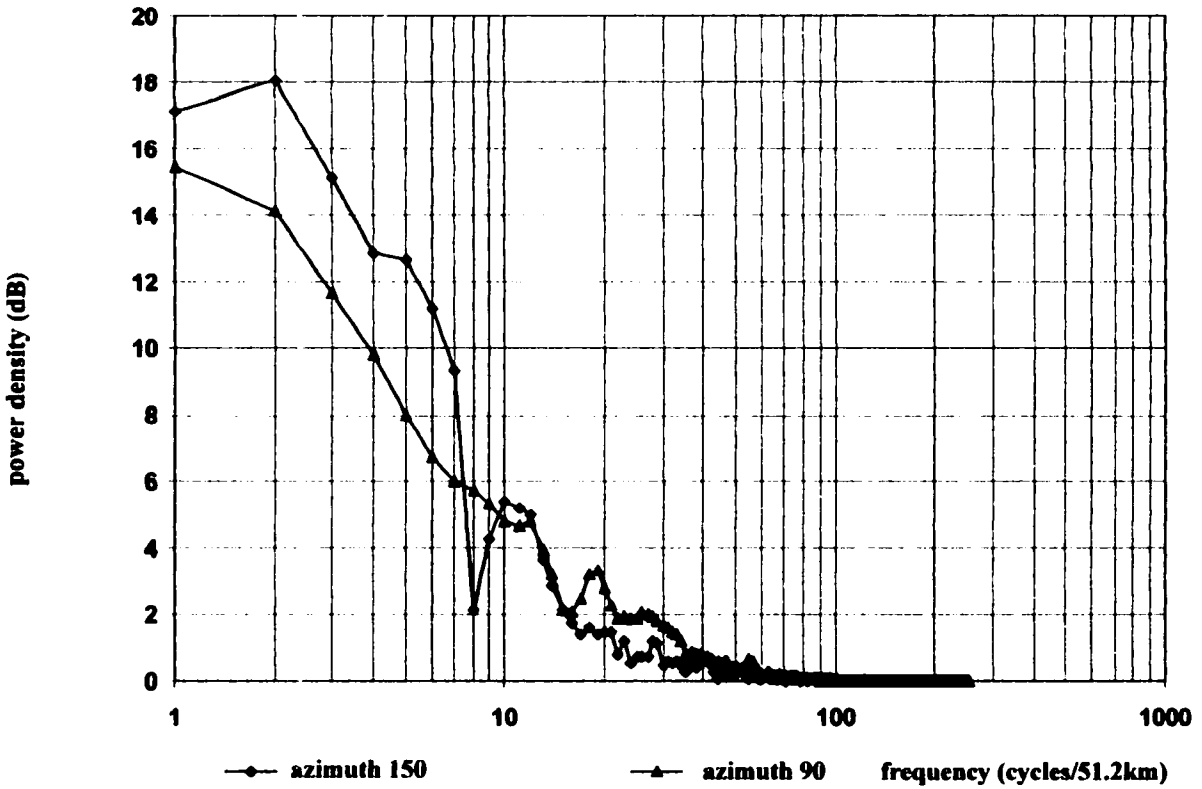


Fig. 6.13: Power spectral density profiles of the terrain surface. X-axis: frequency in logarithmic scale, and unit is cycles/51.2 km; Y-axis: power density in dB.

anisotropy ellipse at the medium and far lag distance. The power spectral profile in the azimuthal direction of 90° shows a number of weak peaks, with the largest peak (relatively) occurring at a frequency of about 19 cycles/51.2 km. This indicates the average frequency of cyclic configuration of small scale features-the aretes (eroded ridges) and cirque glacier troughs in this direction. Half the average wavelength (1.35 km) is close to the autocorrelation length of the minimum axis of the anisotropy ellipse at a small lag distance.

The topography of this area is manifestations of the combination of tectonic and glacial processes acting at different scales. The big glacier valleys and mountain ranges were formed by tectonic forces, while the small scale aretes and cirque depressions were created by the relief-forming activity of local glaciation on mountain slopes. To a large extent, the orientations and shapes of the geomorphologic features condition the anisotropic autocorrelation pattern of DEM errors. The elongated shape of glacier valleys and mountain ranges leads to the high anisotropy ratio of autocorrelation.

Conventionally, the surface slope is used to estimate the error in DEMs as well as in contour maps. However, this study suggests that the linear relationship between the magnitude of DEM error and terrain surface slope is only applicable to a smooth terrain surface (Liu and Jezek, 1999). For a rugged terrain surface, the linear relationship no longer holds, and therefore the surface slope cannot be simply used to predict the error magnitude. The preliminary results from this study also shows that the mean error of a rugged terrain is greater than that of a smooth terrain, and that the elevation values tend to be underestimated in the ridges and overestimated in the valleys.

In summary, the spatial structure of DEM errors is closely related to the terrain surface under investigation. For terrain that contains geomorphologic features of varying scales created by different geophysical processes, the spatial autocorrelation pattern of the DEM error is anisotropic and scale-dependent. The orientation of terrain features determines the direction of maximum and minimum autocorrelation and the wavelength of terrain features controls the autocorrelation range (length). This result has a wide range of implications. In the absence of explicit information on the spatial structure of DEM errors, we still can get qualitative information about the maximum and minimum directions of spatial autocorrelation of the error by identifying the orientation of the dominant terrain features at each scale. The autocorrelation range (length) can be approximately determined by studying the shape and wavelength of the morphological features. If a local interpolator is used to estimate the error at an unsampled point from a finite set of sample points, the shape of the neighborhood search windows should be adapted according to the anisotropic pattern of the error. The sample points that fall in the maximum autocorrelation direction should be given larger weight in the interpolation than those that fall in the perpendicular direction. Since the anisotropic pattern of the error may vary with the spatial separation distance, the error pattern derived at a specific scale may not be applicable at different scales.

The above error analysis results can be used in the accuracy assessment of the terrain parameters derived from our Antarctic DEM and in the design of interpolators for the future improvement of the DEM.

CHAPTER 7

DEM REFINEMENTS BY SPACEBORNE TECHNIQUES

7.1 Improvement of large-scale DEM

By exploiting the best elements of available topographical data, we created the continental scale DEM over the Antarctic. As demonstrated in the preceding chapter, our DEM is superior to previous efforts in terms of spatial resolution and vertical accuracy, as well as reliability and topological consistency. Nevertheless, we recognize that the accuracy and resolution of our DEM greatly depend upon the quality of the topographical source data. In the low slope ice sheet, the DEM is mainly derived from satellite radar altimeter data, and therefore characterized by relatively high vertical accuracy and fairly low spatial resolution. For mountainous areas, cartographic data from the ADD are the primary source data for our DEM. Since high frequency topographical components (small-scale features) were filtered out from the original source data during the contouring process, the DEM derived from cartographic data also has a limited accuracy. The more serious problem with our DEM is that for some regions of the Antarctic no reliable topographical data exist. For example, the Prince Charles Mountains in Lambert glacier basin, some parts of Pensacola Mountains and Shakleton Range, and the flat plateau inside the 81.4°S latitude circle are only covered by very small-scale

topographical maps. Therefore, the accuracy of the DEM is questionable over these regions.

To further improve the Antarctic DEM in the future, we must acquire topographical data of better quality and incorporate them into the DEM. One approach is to gather new large-scale topographic maps. For example, a considerable number of Japanese topographic maps exist at the scale 1:50,000 in the area of Sør Rondane Mountains of the Antarctic. These maps must be digitized for updating the DEM. Another approach is to utilize advanced remote sensing techniques to acquire more reliable and accurate data for improving our DEM, especially in areas where our DEM is of poor quality.

Spaceborne remote sensing techniques represent the most appropriate and efficient means of enhancing and updating a large-scale DEM. In the case of the Antarctic, optical and SAR stereo, InSAR and shape-from-shading techniques can be applied to the existing and future satellite data for improving the DEM quality. During the first Antarctic Imaging Campaign (AIC) in September and October 1997, the Canadian Radarsat-1 captured high resolution SAR image data over the entire Antarctic continent. The combination of images from different beam modes constitutes a large volume of SAR stereo data (Jezek et al, 1998). During the 30 days mission, repeat-pass interferometric SAR data were also produced, which cover a considerable portion of the Antarctic continent (Forster et al, 1998). Moreover, ERS Tandem InSAR data and SPOT optical stereo data also exist for the Antarctic. The Geoscience Laser Altimeter System (GLAS), presently under development, will fly onboard ICESAT with an orbital inclination 94° in the year 2001. With a small footprint of 70 m and 170 m sampling spacing along track, GLAS will provide accurate and dense measurements over the Antarctic up to the latitude

of 86°S. Over most low slope areas of the ice sheets, the vertical accuracy of the elevation measurement is expected to be 15 cm (GLAS Science Team, 1997; Schutz, 1998).

In the following sections, we will report our initial experiments with shape-from-shading, SAR stereo and InSAR techniques over three test sites of Antarctica.

7.2 Shape-from-shading technique

7.2.1 Overview of shape-from-shading technique

Shape-from-shading was originally developed in the field of robot computer vision (Horn, 1990). It is also known as photoclinometry by astrogeologists (Wildey, 1975; Horn, 1990). Its objective is to reconstruct the surface shape and orientation by using image shading information. Shading refers to the continuous variation of image gray tone (brightness or intensity) over space. By introducing constraints and boundary conditions, shape-from-shading can potentially provide a relative change in height at each pixel, leading to a dense set of height measurements and a more faithful rendition of the local terrain shapes.

In essence, shape-from-shading deals with the inverse problem of image formation, involving the inference of surface shape (orientation, slopes) from the spatial variation of image intensity. To derive a numerical solution, an analytical relationship, known as a reflectance map function or photometric function, is often introduced to specify the relationship between image brightness values I(x, y) on the one hand and surface slope, orientation, surface albedo, illumination and imaging sensor vector on the other hand (Frankot and Chellappa, 1989):

$$I(x, y) = R(z_x, z_y, \beta, l, \rho) \tag{7.1}$$

where *I(x, y)* is observed image intensity; *z(x, y)* is unknown surface height above the *(x, y)* plane; $z_x = \partial z / \partial x$ and $z_y = \partial z / \partial y$ are the surface slopes respectively along the x-axis and y-axis; β is the illumination direction vector; *l* is the vector from the surface to the sensor; and ρ is the albedo (intrinsic reflectivity) or normalized backscattering coeffient of the surface cover. Mathematically, the shape-from-shading problem is to find surface orientation (z_x, z_y) or surface height information z(x y) given the image I(x, y) and reflectance map R. To make the reflectance map function tractable, most shape-from-shading algorithms assume that the surface under investigation has a uniform reflecting property, and that illumination direction vector *l* and the viewing direction β are spatially invariant. Therefore, the reflectance map function is simplified as:

$$I(x, y) = R(z_x, z_y) \tag{7.2}$$

It should be noted that even with a known, simplified reflectance map function, the image intensity value at each pixel is a function of the two surface gradient components as shown in equation (7.2). Namely, it has two degrees of freedom at the corresponding location. In its full generality, the shape-from-shading problem is an inherent underdetermined problem (ill posed problem). To derive a shape-from-shading algorithm, either additional constraints, a priori knowledge, or simplifying assumptions need to be introduced.

Permanent ice and snow cover the overwhelming majority of the Antarctic continent. Previous research has shown that satellite images encode the subtle variation of topographic surface in the Polar Regions (Lodwick and Paine, 1985; Dowdswell and

McIntyre, 1987; Bindshadler and Vornberger, 1994; Scambos and Fahnestock, 1998). In comparison with most terrestrial surfaces, the ice and snow covered Antarctic has much more uniform reflecting characteristics (albedo or scattering cross-section coefficient). According to our calculation, 97.5% of the Antarctic has surface slopes of less than 3°, therefore most optical images are not affected by terrain-induced shadow although summer solar elevation angle is quite low in the Antarctic. Fairly homogenous reflectivity (albedo) and a smooth and shadow-free surface make the shape-from-shading technique most suitable for the extraction of the detailed shape information about small-scale terrain features over the Antarctic ice sheet. A large volume of SPOT and Landsat images are available up to 82°S. According to Swithinbank (1988), 55% of the continent is covered by good quality cloud-free Landsat images.

7.2.2 Algorithm implementations with application to SPOT image

Two shape-from-shading algorithms were implemented using the C programming language. One is the local closed form algorithm first proposed by Pentland (1990), and the other is the global iterative algorithm with integrability constraint developed by Frankot and Chellappa (1987; 1989). The algorithms were tested on a SPOT HRV panchromatic image with 10 m resolution over the Crary Ice Rise in the Ross Ice Shelf (Fig. 7.1a). The image was acquired by SPOT 3 on Nov. 24, 1995. The center of the scene is 82°59′53″S and 172°02′47″W, covering approximately 60x60 km area. At the image acquisition time, the sun elevation angle is 19.3°, and sun azimuth angle is 99.2°.

The local algorithm is based on the linear approximation of the true reflectance function in terms of the surface gradient (p, q) (Pentland, 1990; Cooper, 1994). When the

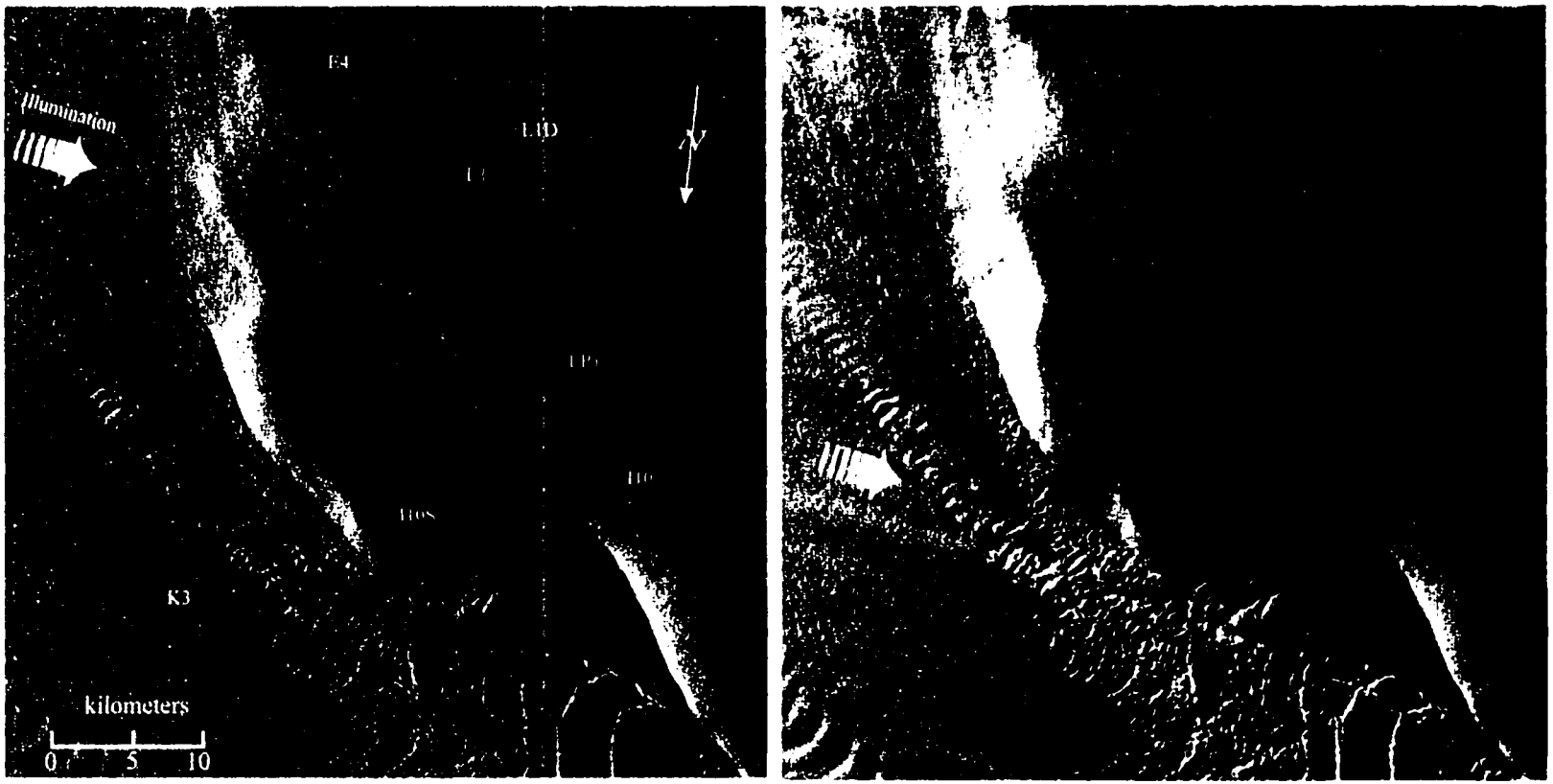


Fig. 7.1: SPOT image and shape-from-shading derived DEM over Crary Ice Rise, Antarctica. (a) SPOT HRV Panchromatic image; (b) hill-shaded elevation data derived from local closed-form shape-from-shading technique.

snow is fresh and free of surface frost, a reflectance function can be approximated by the Lambertian model (Warren, 1982; Scambos and Fahnestock, 1998):

$$I(x,y) = R(p,q) = \frac{ps_x + qs_y + s_z}{\sqrt{p^2 + q^2 + 1}} \tag{7.3}$$

where I(x,y) is the observed gray values of the image; (s_x, s_y, s_z) is the unit illumination direction vector; and $p = \partial z / \partial x$ and $q = \partial z / \partial y$ are the surface slopes along the x and y directions. By taking the Taylor series expansion of the reflectance function, we can form a linear approximation to equation (7.3) around $(p,q) = (p_0, q_0)$:

$$I(x,y) \approx R(p_0,q_0) + (p - p_0)\frac{\partial R}{\partial p}(p_0,q_0) + (q - q_0)\frac{\partial R}{\partial q}(p_0,q_0) \tag{7.4}$$

We obtain the linear approximation of equation (7.3) using formula (7.4) at $p_0 = q_0 = 0$,

$$I(x,y) = ps_x + qs_y + s_z \tag{7.5}$$

By taking the Fourier transform of both sides of equation (7.5) and droping the DC term (s_z) on the right, we get:

$$F_I = s_x F_p + s_y F_q \tag{7.6}$$

Using the theorem of Fourier transform of the derivative (Weaver, 1983), we get:

$$\begin{aligned} F_p &= F\{\frac{\partial Z}{\partial x}\} = 2\pi j \omega_x F_z(\omega_x, \omega_y) \\ F_q &= F\{\frac{\partial Z}{\partial y}\} = 2\pi j \omega_y F_z(\omega_x, \omega_y) \end{aligned} \tag{7.7}$$

Equation (7.6) can be written as:

$$F_I = 2\pi j(\omega_x s_x + \omega_y s_y) F_z(\omega_x, \omega_y) \tag{7.8}$$

$$F_Z(\omega_x, \omega_y) = [2\pi j(\omega_x s_x + \omega_y s_y)]^{-1} F_I \tag{7.9}$$

where F_I is the Fourier transform of the image I(x,y); $F_z(\omega_x,\omega_y)$ is the Fourier transform of surface height Z(x,y); (ω_x,ω_y) are the frequency indices; F_p is the Fourier transform of surface derivative along x direction, F_q is the Fourier transform of surface derivative along y direction; and $j=\sqrt{-1}$ is the imaginary number.

The surface height information Z(x,y) can be derived by taking the inverse Fourier transform of equation (7.9). This local algorithm does not require the explicit reflectance map of the surface. As long as the surface gradients (p, q) are restricted to some range, this Discrete Fourier Transform (DFT) based algorithm can efficiently estimate the surface shape by the use of a known illumination position.

Using the sun elevation and azimuth angles, we first calculated the unit illumination vector and then applied this local algorithm to the SPOT image. The height information derived from this algorithm is arbitrarily scaled. We calibrated a surface by using the Singular Value Decomposition (SVD) algorithm (Press et al, 1992):

$$h = a + bx + cy + dz \tag{7.10}$$

where h is the calibrated absolute elevation value; x and y are the image coordinates; z is the relative height value derived from shape-from shading; and a, b, c, and d are coefficients calibrated. Four points with known geolocation and elevation from the Siple Coast Project (SCP) (Bindschadler et al, 1988) and two points from the RIGGS project (Bentley and Jezek, 1981) are used to calibrate the surface equation (7.10). The locations of these points are shown in Fig. 7.1a. The RMSE of the surface fitting is 3.7 m. The resulting elevation data are represented as a hill-shaded image in Fig. 7.1b and contour map and 3D perspective view in Fig. 7.2. Comparison of the hill-shaded image with the

original SPOT image (Fig.7.1) suggests that the elevation data derived from the shape-from-shading are remarkably consistent with the visual intuition perceived from the real image. The pattern and values of contours (Fig. 7.2) derived from the shape-from-shading agree with the coarse contour map in Bindschadler et al (1990).

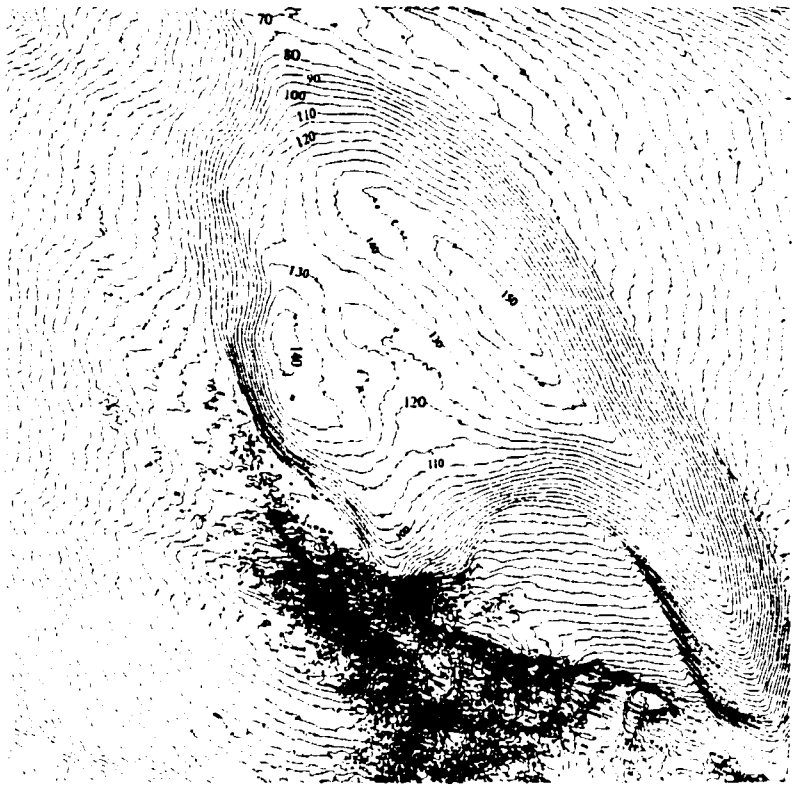
Two types of problems are identified with this local algorithm. First, the Fourier components perpendicular to the illumination direction cannot be seen in the image data, and are simply set to zero to avoid the data overflow of equation (7.8) caused by $\omega_x s_x + \omega_y s_y = 0$. Consequently, some artificial linear features aligned with solar illumination direction appeared in the hill-shaded image (Fig. 7.1b). The other problem is that many ice crevasses are distributed in the lower right part of the 'finger' of the ice rise. These rough, crumpled, and wrinkled ice crevasses have different and varying reflectivity and violate the uniform albedo assumption. As shown in Fig. 7.2, we obtained erroneous elevation values in that part.

Also, a global shape-from-shading algorithm with the integrability constraint was implemented in this research. This algorithm was originally derived by Brooks and Horn (1985) by employing variational calculus. Frankot and Chellappa (1987) introduced the strict integrability constraint so that it converges faster and with much less error than the original version. The algorithm seeks the surface slope estimates $(\hat{z}_x, \hat{z}_y)$ that minimize the following cost function:

$$\varepsilon = \iint (I(x,y) - R(\hat{z}_x, \hat{z}_y))^2 + \lambda(\hat{z}_{xx}^2 + 2\hat{z}_{xy}^2 + \hat{z}_{yy}^2) dxdy \qquad (7.11)$$

where $I(x,y)$ is the actual image brightness value; $R(\hat{z}_x, \hat{z}_y)$ is the predicted brightness value; $\hat{z}_x, \hat{z}_y$ are the estimated surface slopes along x and y direction; $\hat{z}_{xx}, \hat{z}_{xy}, \hat{z}_{yy}$ are

(a)



(b)

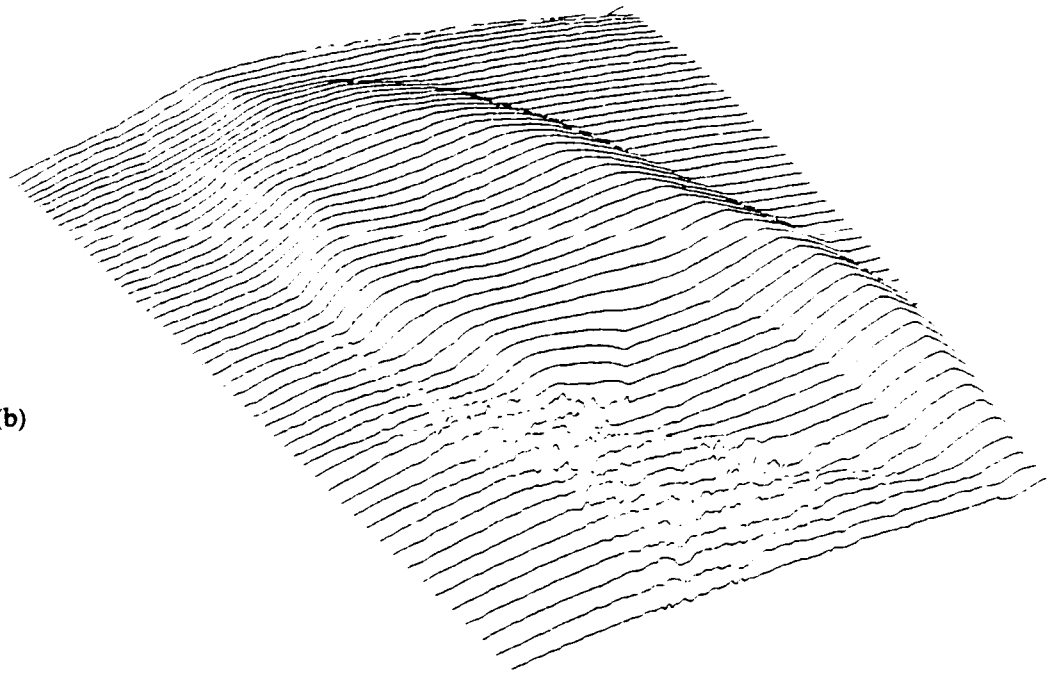


Fig. 7.2: Elevation data derived from local shape-from-shading technique. (a) contours with an interval of 5 m; (b) 3D perspective view.

the second partial derivative of Z(x,y); and λ is the regularization parameter.

The first term in equation (7.11) is the brightness constraint. It requires that the reconstructed shape produce the same brightness as the input image at each surface point. The second term is a measure of the quadratic variation of surface slopes, forcing the gradient of the surface to change smoothly. The inclusion of this smoothness constraint regularizes the original ill-posed shape-from-shading problem in the form of a regularization penalty function and assures a unique smooth solution to equation (7.11). The regularization parameter λ sets the tradeoff between the relative contributions of the brightness error term and the term measuring departure from smoothness (allowable surface oscillation).

An iterative solution was derived by Brooks and Horn (1985) using the Euler-Lagrange equations to minimize equation (7.11) subject to the constraint that $(\hat{z}_x, \hat{z}_y)$ satisfy known boundary conditions. The Jacobi iterative scheme is used, and a smoothing step is used between successive Jacobi iterations to stablize the iteration process. However, like most other global shape-from-shading algorithms, Brooks and Horn's solution is not integrable. Integrability is defined as

$$\frac{\partial^2 Z}{\partial x \partial y} = \frac{\partial^2 Z}{\partial y \partial x} \tag{7.12}$$

The integrability constraint requires the second partial derivatives be independent of the order of differentiation and interrelates the two components of surface orientation $(\hat{z}_x, \hat{z}_y)$, rather than allowing them to vary independently. Consequently, the surface elevation recovered at any particular point is independent of the path of integration. In practice, enforcing this condition also smoothes the computed terrain surface. Frankot

and Chellappa (1987; 1989) added the strict integrability constraint in Brooks and Horn's iterative algorithm. The basic idea is to project the possibly nonintegrable surface slope estimates at each iteration step onto the nearest integrable surface slopes by the DFT transformation:

$$\tilde{z}(\omega) = \frac{a_x^*(\omega) \cdot \hat{z}_x(\omega) + a_y^*(\omega) \cdot \hat{z}_y(\omega)}{|a_x(\omega)|^2 + |a_y(\omega)|^2} \tag{7.13}$$

where $\tilde{z}(\omega)$ is the Fourier transform of the integrable height $z(x,y)$; $\hat{z}_x(\omega)$ and $\hat{z}_y(\omega)$ are the Fourier transform of the newest estimated slopes in the x-direction and y-direction; $a_x(\omega)$ and $a_y(\omega)$ are Fourier coefficients of a discrete differentiation operator in x and y; $a_x^*(\omega)$, $a_y^*(\omega)$ are the conjugates of $a_x(\omega)$ and $a_y(\omega)$; and $\omega = (\omega_x, \omega_y)$ is the 2 dimensional Fourier domain frequency coordinates. The discrete Fourier basis projection in the equation (7.13) makes the projected surface slopes satisfy the integrability condition of equation (7.12), while simultaneously minimizing the distance between the old surface slope estimates and the projected, integrable surface slopes:

$$d\{(\hat{z}_x, \hat{z}_y), (\tilde{z}_x, \tilde{z}_y)\} = \iint (|\tilde{z}_x - \hat{z}_x|^2 + |\tilde{z}_x - \hat{z}_x|^2) dx dy \tag{7.14}$$

The equation (7.13) is valid except at the point $\omega = (0,0)$, which implies that we cannot recover the average value of Z (DC component).

The iterative processing procedure can be summarized as the following steps:

(1) Given the previous slope estimates $(\hat{z}_x, \hat{z}_y)$, perform the smoothing operation to stablize the Jacobi iteration.

$$\hat{\bar{z}}_x(n,m) = \frac{1}{5}[\hat{z}_x(n,m+1) + \hat{z}_x(n,m-1) + \hat{z}_x(n+1,m) + \hat{z}_x(n-1,m)] + \frac{1}{20}[\hat{z}_x(n-1,m-1) + \hat{z}_x(n-1,m+1) + \hat{z}_x(n+1,m+1) + \hat{z}_x(n+1,m+1) \tag{7.15}$$

Similarly for $\hat{\bar{z}}_y(n,m)$;

(2) Calculate the image gray value $R(\hat{\bar{z}}_x, \hat{\bar{z}}_y)$ at every pixel (n,m) by substituting the smoothed surface slope estimates into the reflectance map formula R;

(3) Calculate partial derivatives of $R(\hat{\bar{z}}_x, \hat{\bar{z}}_y)$ with respect to z_x and z_y at every pixel;

$$\left.\begin{matrix} \dfrac{\partial R(z_x, z_y)}{\partial z_x}\Big|_{\substack{z_x = \hat{\bar{z}}_x \\ z_y = \hat{\bar{z}}_y}} \\ \dfrac{\partial R(z_x, z_y)}{\partial z_x}\Big|_{\substack{z_x = \hat{\bar{z}}_x \\ z_y = \hat{\bar{z}}_y}} \end{matrix}\right. \tag{7.16}$$

(4) Update the slope estimates using the Jacobi recursion from (k)th iteration to (k+1)th iteration;

$$\begin{bmatrix} \bar{z}_x \\ \bar{z}_y \end{bmatrix}_{k+1} = \begin{bmatrix} \hat{\bar{z}}_x \\ \hat{\bar{z}}_y \end{bmatrix}_k + \frac{3}{10\lambda}(I(x,y) - R(\hat{\bar{z}}_x, \hat{\bar{z}}_y)) \begin{bmatrix} \dfrac{\partial R(\hat{\bar{z}}_x, \hat{\bar{z}}_y)}{\partial z_x} \\ \dfrac{\partial R(\hat{\bar{z}}_x, \hat{\bar{z}}_y)}{\partial z_y} \end{bmatrix}_k \tag{7.17}$$

(5) Perform the discrete Fourier transform (DFT) of $\bar{z}_x, \bar{z}_y$, obtaining $\hat{z}_x(\omega), \hat{z}_y(\omega)$.

(6) Substitute $\hat{z}_x(\omega), \hat{z}_y(\omega)$ into equation (7.13) to enforce integrability, obtaining the Fourier transform of integrable surface elevation $\bar{z}(\omega)$ and Fourier transforms of the nearest integrable surface slopes $\bar{z}_x(\omega)$, $\bar{z}_y(\omega)$.

$$\tilde{z}_x(\omega) = a_x(\omega) \cdot \tilde{z}(\omega)$$

$$\tilde{z}_y(\omega) = a_y(\omega) \cdot \tilde{z}(\omega)$$

$$a_x(\omega) = j\sin(\omega_x)$$

$$a_y(\omega) = j\sin(\omega_y)$$

$$j = \sqrt{-1}$$

$$\omega = (\omega_x, \omega_y) \in (\frac{2\pi n}{N}, \frac{2\pi m}{N})$$

$$n, m \in (0,1,\ldots, N-1\}$$

(7) Perform inverse DFT of $\tilde{z}_x(\omega), \tilde{z}_y(\omega)$ to obtain the newest integrable surface slopes; Repeat steps (1)-(6) until the cost function in (7.11) becomes sufficiently small or stops decreasing. The final surface height is obtained by the inverse Fourier transform of $\tilde{z}(\omega)$.

This global algorithm was tested on the SPOT image of Crary Ice Rise. The initial surface slopes were set to zero and λ=7000. Fig. 7.3 shows the results after 5, 30, 100 and 500 iterations. The RMSE values of surface fitting with the six known points are respectively 7 m, 6 m, 5 m, and 2.8 m. Visual inspection of the surfaces in Fig. 7.3 and the decrease in the RMSE of surface fitting suggest that the quality of the surface derived from the global shape-from-shading method is improved with the increase of iterations. Fig. 7.4 shows the contours and perspective view of the surface derived at the iteration of 1000. The RMSE of surface fitting with the known points is only 2.5 m. In comparison with the local method, this global method is more robust with respect to local noise. The contours derived from the global method are much smoother than the local method. The magnitude of the errors induced by the ice crevasses is also much less than the local

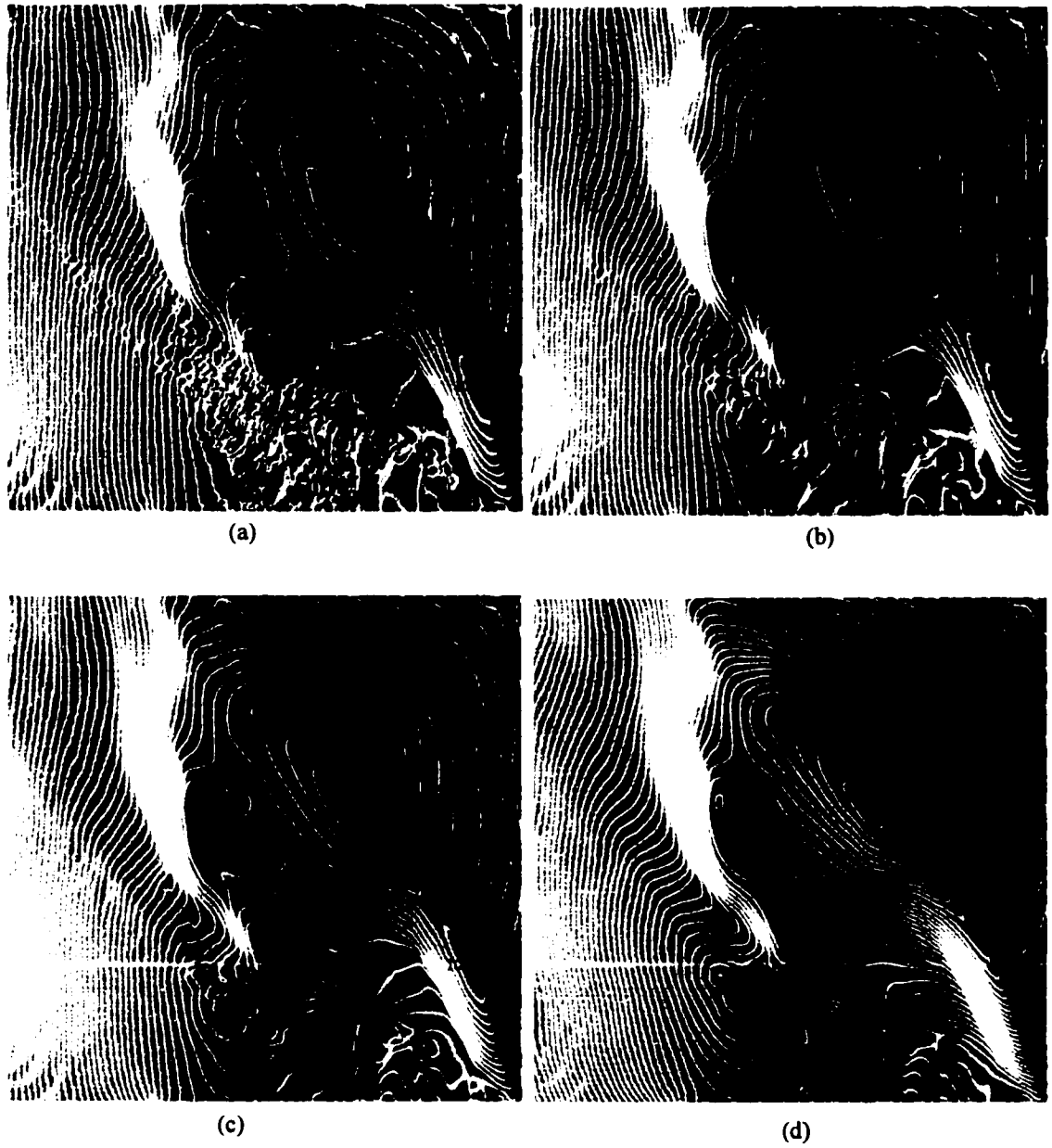


Fig. 7.3: Elevation data derived from global shape-from-shading technique. (a) iterations = 5; (b) iterations = 30; (c) iterations = 100; (d) iterations = 500

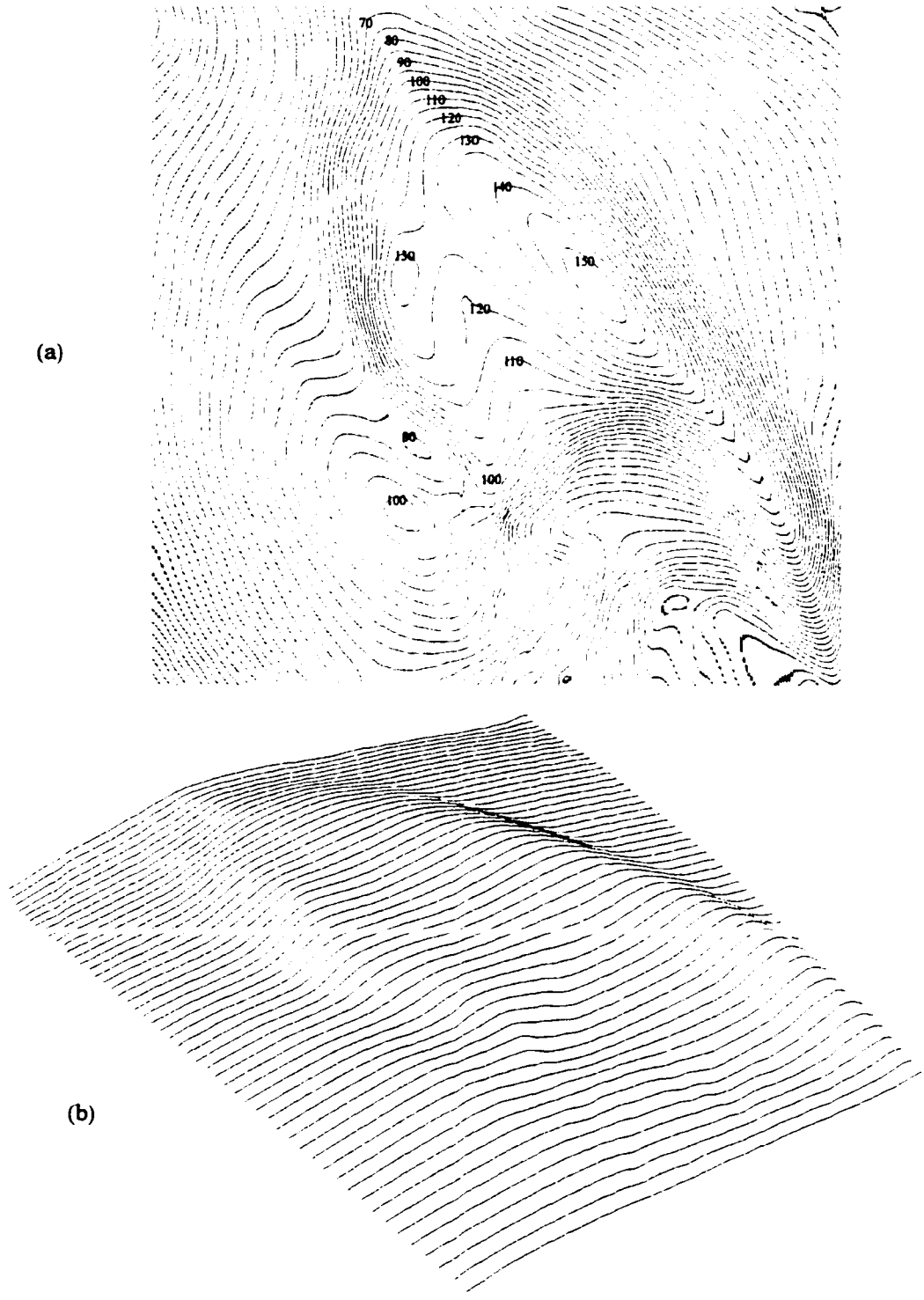


Fig. 7.4: Contours and perspective view of DEM derived from global shape-from-shading algorithm (iterations = 1000). (a) contours with an interval of 5 m; (b) 3D perspective view.

method. Nevertheless, the global method is computationally more costly.

In summary, our shape-from-shading experiments demonstrate that optical satellite images can be utilized to enhance the spatial resolution of our DEM over the low-slope ice sheet of the Antarctic. With the shape-from-shading and a number of known elevation points, a dense elevation grid with 2~4 m of accuracy are achievable. Our experiment results also suggest that if the surface is smooth and gently undulated, the local method is sufficient for deriving a reliable DEM. If the surface is noisy and contains ice crevasses, the global method should be used because it is more robust and reliable. In addition, it is found that the elevation data derived from the local shape-from-shading method can be used as initial seed values for the global iterative shape-from-shading to speed up the convergence of the iteration process.

Previous experiments with SAR images over other areas also suggest that the shape-from-shading method is applicable to SAR imagery (Frankot and Chellappa, 1987; Thomas et al., 1991). In addition to ERS-1, ERS-2, and JERS-1 imagery, Radarsat SAR images acquired during the Antarctic Imaging Campaign (AIC) cover the complete continent, providing a potential image source for SAR-based shape-from-shading application.

7.3 SAR stereo technique

7.3.1 Principles of SAR stereo technique

The SAR stereo technique derives the digital elevation data based on parallax differences between an stereo image pair. A stereo pair can be formed by two images of the same area acquired from two different viewpoints along adjacent satellite orbits. Though

similar to the optical stereo photogrammetric technique in principle, the SAR stereo technique is based on an entirely different projection geometry and mathematical models.

There are two possible SAR stereo geometric arrangements: same-side and opposite-side. As SAR actively illuminates the object, a difference in imaging geometry also implies a difference in illumination. Larger angular separation between two images of a stereo pair tend to provide a more stable solution from a geometric point of view, but may induce excessive radiometric differences between two images, causing difficulties in identifying conjugate points between images. Previous experiments show that stereo correlation (viewability) is ensured at shallow look angles for same-side arrangements, but opposite-side stereo is feasible only with flat or gently rolling terrain (Leberl, 1990).

Mathematically, the SAR stereo technique attempts to calculate the three-dimensional position of a ground target point P(X,Y,Z) given its image coordinates (i,j) and (i',j') in two input stereo images, master and slave. For the image coordinate pair (i,j) in the master image, we can form two equations in a geocentric Cartesian XYZ system: range equation (7.18) and Doppler equation (7.19):

$$R_s = \sqrt{(\vec{S} - \vec{P}) \cdot (\vec{S} - \vec{P})} \tag{7.18}$$

$$f_D = \frac{2}{\lambda} \frac{(\vec{V}_S - \vec{V}_P) \cdot (\vec{S} - \vec{P})}{R_s} \tag{7.19}$$

where: $\vec{P} = (X, Y, Z)$: unknown position of the ground point to be solved;

$\vec{S} = (X_S, Y_S, Z_S)$: sensor position at the time the ground point is imaged;

$\vec{V}_P = (\dot{X}, \dot{Y}, \dot{Z})$: velocity vector of the ground point;

$\vec{V}_S = (\dot{X}_S, \dot{Y}_S, \dot{Z}_S)$: velocity vector of the sensor along the orbit;

R_s : slant range from the sensor to the ground point;

λ : radar wavelength;

f_D : the Doppler frequency shift.

For squint angle τ, the angle between the line connecting the sensor/ground target and the zero Doppler plane, the Doppler centroid is given by:

$$f_{DC} = \frac{2\left|\vec{V}_s - \vec{V}_P\right|}{\lambda}\sin\tau \tag{7.20}$$

If the squint angle τ equals zero, the Doppler centroid f_{DC} is the zero Doppler. Otherwise, f_{DC} varies over the range and can be approximated by fitting a polynomial as a smooth function of the range coordinate. As a SAR sensor continuously moves along an orbit during image acquisition, the sensor state vector $\vec{S} = (X_S, Y_S, Z_S)$ and velocity vector $\vec{V}_S = (\dot{X}_S, \dot{Y}_S, \dot{Z}_S)$ can be modeled by low order polynomials using satellite ephemeris data, and can be further refined by using the ground control points (GCPs).

Given the image coordinates (i,j) along with satellite ephemeris data, we can determine the Doppler frequency shift f_{DC}, slant range R_s and imaging time t (interior orientation) (Fig. 7.5). Substituting imaging time into the fitted orbital model, we can obtain the sensor state vector $\vec{S} = (X_S, Y_S, Z_S)$ and velocity vector $\vec{V}_S = (\dot{X}_S, \dot{Y}_S, \dot{Z}_S)$ (exterior orientation parameters) at the time the image pixel (i,j) was at the center of the radar beam. Therefore, the range equation (7.18) and the Doppler equation (7.19) only contain three unknowns $\vec{P} = (X, Y, Z)$, namely, the geolocation of the ground point associated with the image pixel (i,j).

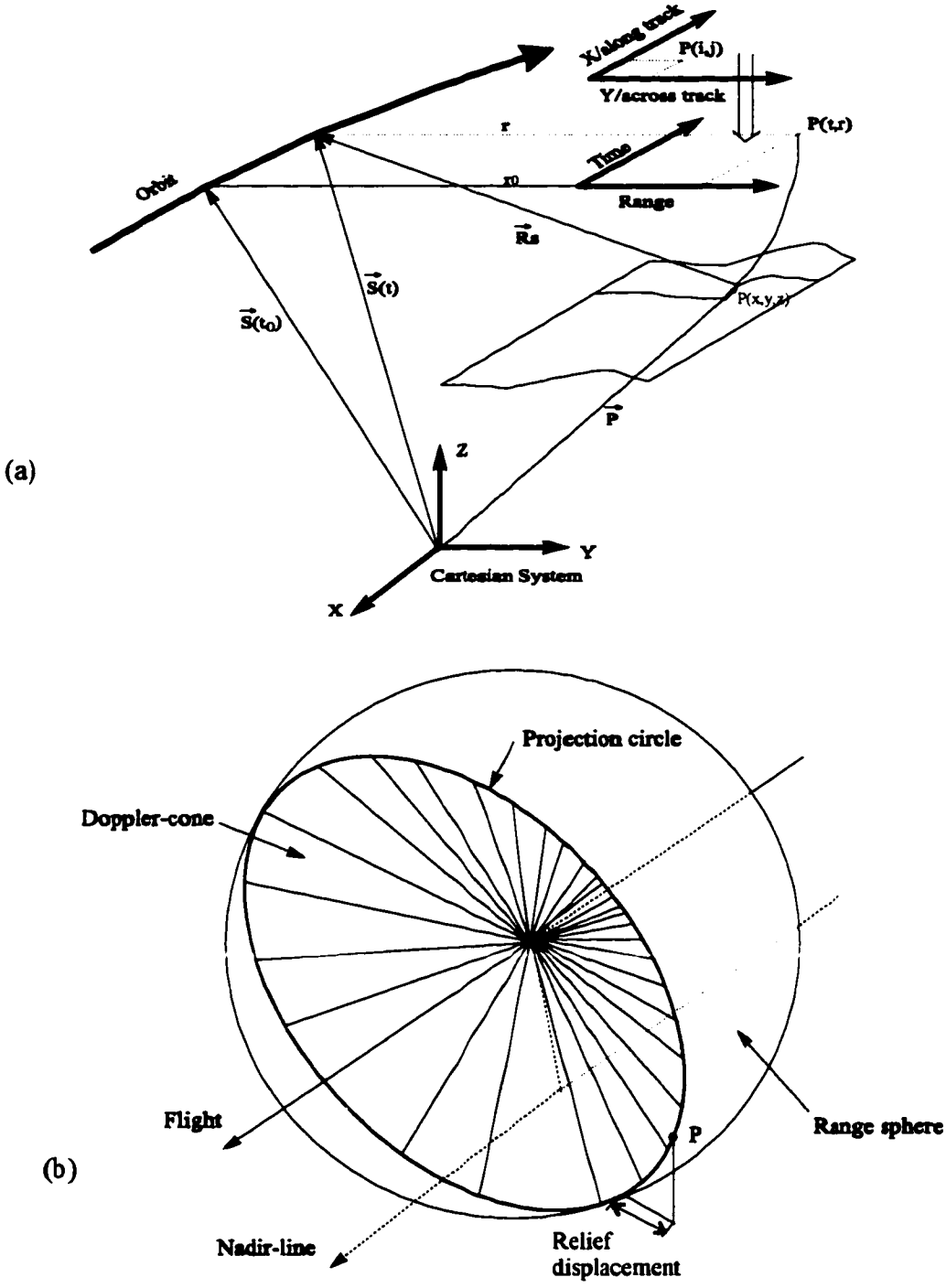


Fig. 7.5: Sensor state vector determination and range/Doppler intersection in SAR stereo techniques. (a) image to object coordinate system conversion technique; (b) intersection of range sphere and doppler cone.

The range equation (7.18) determines a range sphere with radius R_s, saying that the ground target point must lie on this sphere centered at the sensor. The Doppler equation (7.19) describes a Doppler cone whose vertex is at the sensor (Fig. 7.5), saying that the target point must be located along the cone. When the squint angle is zero, the Doppler cone is fully opened and degenerates to a plane. The intersection of the range sphere and the Doppler cone is a circle centered at the sensor flight line. Obviously, with one SAR image the geolocation of the image pixel (i,j) can only be unambiguously determined to the locus of points on the range/Doppler circle rather than a single point.

Similarly, using the conjugate image pixel (i',j') in the slave image, we can obtain two other equations:

$$R_s' = \sqrt{(\vec{S}' - \vec{P}) \cdot (\vec{S}' - \vec{P})} \tag{7.21}$$

$$f_D' = \frac{2}{\lambda} \frac{(\vec{V}_S' - \vec{V}_P) \cdot (\vec{S}' - \vec{P})}{R_s'} \tag{7.22}$$

Having the four equations of (7.18), (7.19), (7.21) and (7.22) in hand, the three unknowns $\vec{P} = (X, Y, Z)$ are overdetermined. In practice, we can first intersect the two range spheres associated with a conjugate pixel pair to get the range circle. The range circle can be intersected with one of the range/Doppler circles. The intersection point farthest from the Earth center can be trivially discarded based on knowledge of the imaging geometry and object space coordinate system. The four equations can be also iteratively solved using a least square adjustment approach.

The accuracy of the elevation data derived from SAR stereo is influenced by the measurement error of the radar slant range, mismatch of conjugate image pixels, and

errors of the sensor position and attitude (velocity vector). The basic limitation is the measurement error of radar slant range (Leberl, 1998):

$$\sigma_h^2 = \frac{\sin^2\theta - \sin^2\theta'}{\sin^2(\theta - \theta')} * \sigma_r^2 \qquad (7.23)$$

where σ_h is the error of derived terrain height; σ_r is the measurement error of slant range; and θ and θ' are looking angles of SAR sensor used to acquire the two stereo images (master and slave).

7.3.2 SAR stereo processing with Radarsat data

During the first Antarctic Imaging Campaign (AIC) in September and October 1997, the Canadian Radarsat-1 acquired high resolution SAR image data over the entire Antarctic continent. Radarsat has the unique capability of collecting image data with a number of beam pointing modes. The Standard Beam 2 was selected for the Radarsat Antarctic Mapping Project (RAMP) as the primary imaging mode for optimizing ice sheet observations, and other higher incidence angle beams were also used during the 30 days of image acquisition period, including the Standard Beam 7, Standard Beam 3, 4, 5, 6, Extended High-Beam 4, and ScanSAR. The combination of different beam modes constitutes a large volume of SAR stereo data.

SAR stereo analysis was performed over two test sites, the Terra Nova Bay area and the Mackey Glacier area of the Transantarctic Mountains. Here we only report the result for the Terra Nova Bay area. The Standard beam 2 (ST2) image (Fig. 7.6a) was acquired on Oct 9, 1997 with an incident angle of 28.02° at the scene center, and Standard beam 7 (ST7) image (Fig. 7.6b) on Sep. 20, 1997 with an incident angle of 47.22° at the scene

center, resulting in about a 19.2° intersection angle. As shown in Fig. 7.6, the terrain is characterized by glacial landforms. The rugged mountain slopes, relatively flat glacial valleys, and the volcano-Mt. Melbourne can be observed in the scene. Stereo processing is performed by using the Vexcel Radar Stereo Toolkit-RaST™ (Marra et al, 1998).

For the stereo pair, we conducted a series of stereo processing steps, including image speckle reduction, coarse image co-registration, hierarchical image matching, sensor state vector determination and range/Doppler intersection, error detection and removal, and spatial interpolation.

Two original stereo image files are in SGF (SAR Georeferenced Full Resolution Product) format with 16 bit gray values. They are first processed to extract the meta data, which contain the information about sensor position, attitude and imaging geometry. The 16 bit images are simply normalized to 8 bit gray scale images. Then, the edge preserving filters, Lee's sigma filter, is applied to the 8 bit images to remove the radar speckles to improve the quality of the subsequent image matching process.

To support efficient and accurate matching, the ST2 image was resampled with reference to the ST7 image through an affine transformation. Nine evenly distributed tie points were selected between the two images using an on-screen manual cursor-pointing method, and derived the affine transformation coefficients. This transformation establishes a coarse co-registration between images and removes scale, translation and relative rotation differences. Prior to range Doppler intersection, an inverse affine transformation was applied to the resulting match points.

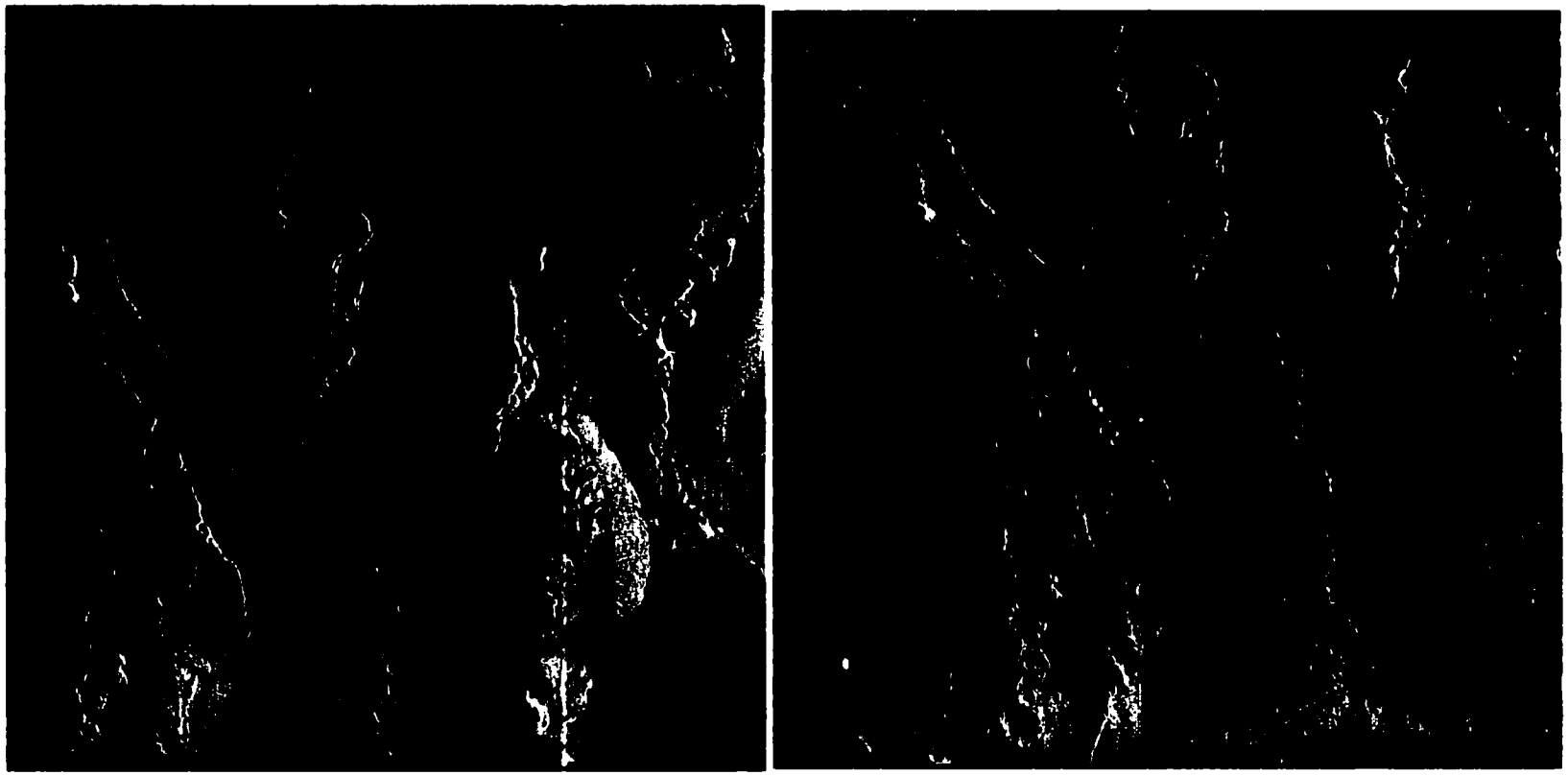


Fig. 7.6: Stereo pair of RADARSAT SAR data over Terra Nova Bay area
(a) standard beam 2 image; (b) standard beam 7 image.

Identifying a dense set of conjugate pixels between the two images is the most important step in SAR stereo processing. A hierarchical area-based method was used for the initial matching. This hierarchical matching processing continued for eight levels for this image pair. To avoid poor and false matches, the matching quality was controlled by a correlation peak value, and correlation peak on edge of search space. A match candidate is rejected if either of two indicators is below its threshold value.

Then, the epipolar geometry was used to remove the parallaxes in one direction. This changes the two-dimensional conjugate pixel searching into one dimensional searching problem. The use of epipolar geometry makes the image matching process more efficient. The matching results from the two-dimensional hierarchical matching were refined by the epipolar matching process.

At the final stage, we totally obtained 1,857,697 pairs of matched pixels, and the average density is about 186 points/km^2. As shown in Fig. 7.7, matching points are generally very dense. However, data holes are observed in the areas of flat glacier floors, and high relief areas. In the glacial floors, images are devoid of distinct features and characterized by a periodic or repetitive textual pattern, leading to poor matching results. Layover and excessive foreshortening generated considerable geometric distortion and radiometric difference between the two stereo images and failed the automated matching. In addition, areas with poor image texture and low signal-noise ratio also lead to poor matching results.

Nine Ground Control Points (GCPs) were manually identified by using 1:250,000 scale Italian and USGS topographic map and GPS measurements over the Terra Nova Bay area. Prior to stereo intersection computation, the position and attitude of the sensor

Fig. 7.7: Distribution of matched points from SAR stereo processing

were refined by orbit adjustment.

For each conjugate pixel pair in the dense set of matched points, the software automatically calculated the slant range, imaging time, and position and velocity of the SAR sensor using the GCP-refined orbital model. Based on these calculations, four range and Doppler equations are formed to perform the range/Doppler intersection and point associated with each pair of matched conjugate points.

Despite the quality control at the image matching step, the elevation data directly extracted using the Vexcel SAR Stereo tool are very noisy, and a significant number of erroneous values exist as shown in Fig. 7.8.

To filter out the erroneous values and noisy measurements, error detection and elimination operations were performed. First, a median filter with a 100 m window size was used to reduce the data volume. Then, the whole data set was subdivided into four sections, and each section was checked by our locally adaptive and robust error detection method, which was described in detail in Chapter 3. Many noisy and erroneous data points were successfully detected and filtered out. The resulting surface is presented in Fig. 7.9a. The comparison between Fig. 7.8 and Fig. 7.9a shows the sharpness and effectiveness of our local error detection method.

Fig. 7.9 and Fig. 7.10 compare hill-shaded relief images and contours derived from the SAR stereo technique with those from 1:250,000 USGS topographic map. It is quite clear that the elevation data derived from the SAR stereo technique contain more topographic details than the 1:250,000 USGS topographic map, except for the areas where the original SAR stereo measurements have data gaps (Fig. 7.7). Since no other credible ground truth data are available, we manually extracted 15 spot elevation points

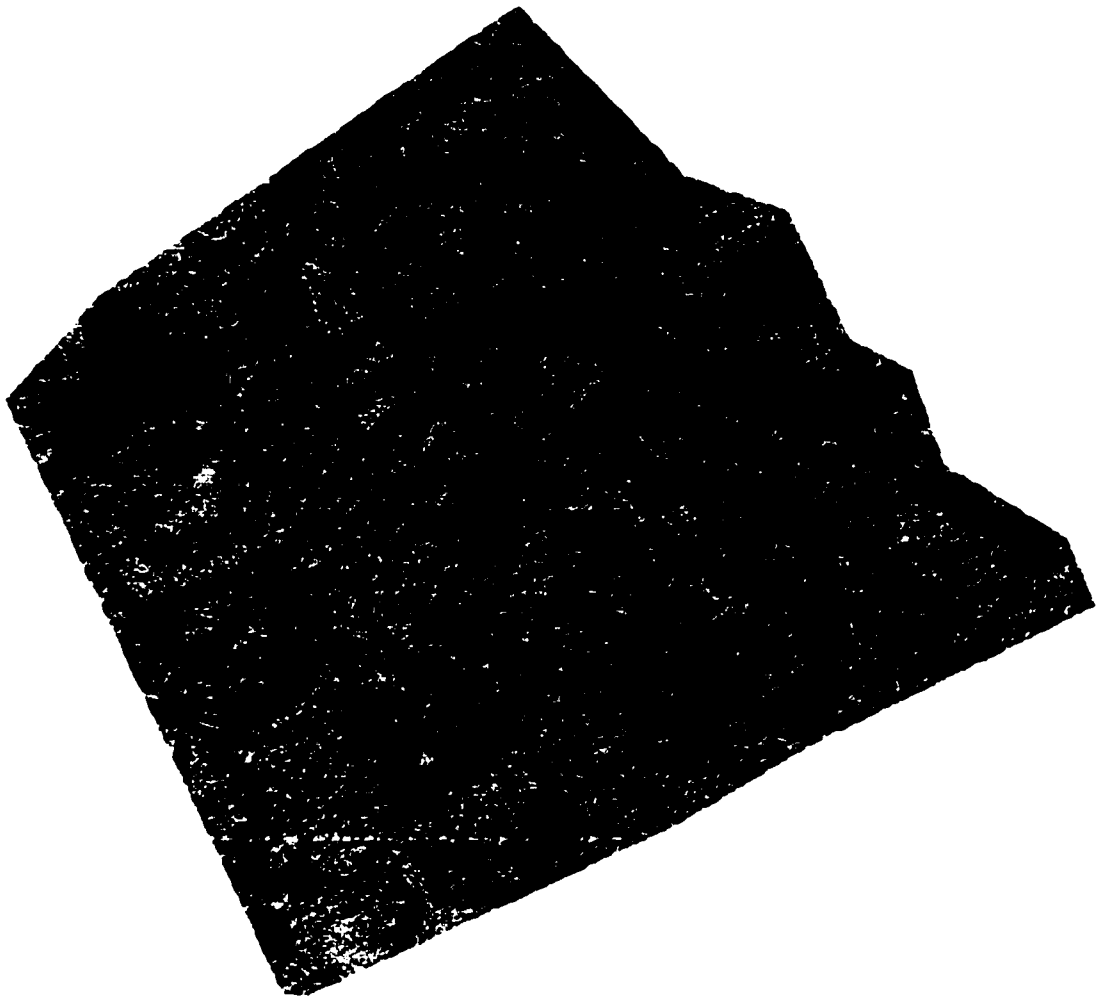


Fig. 7.8 Noisy surface directly derived from the Vexcel stereo tool

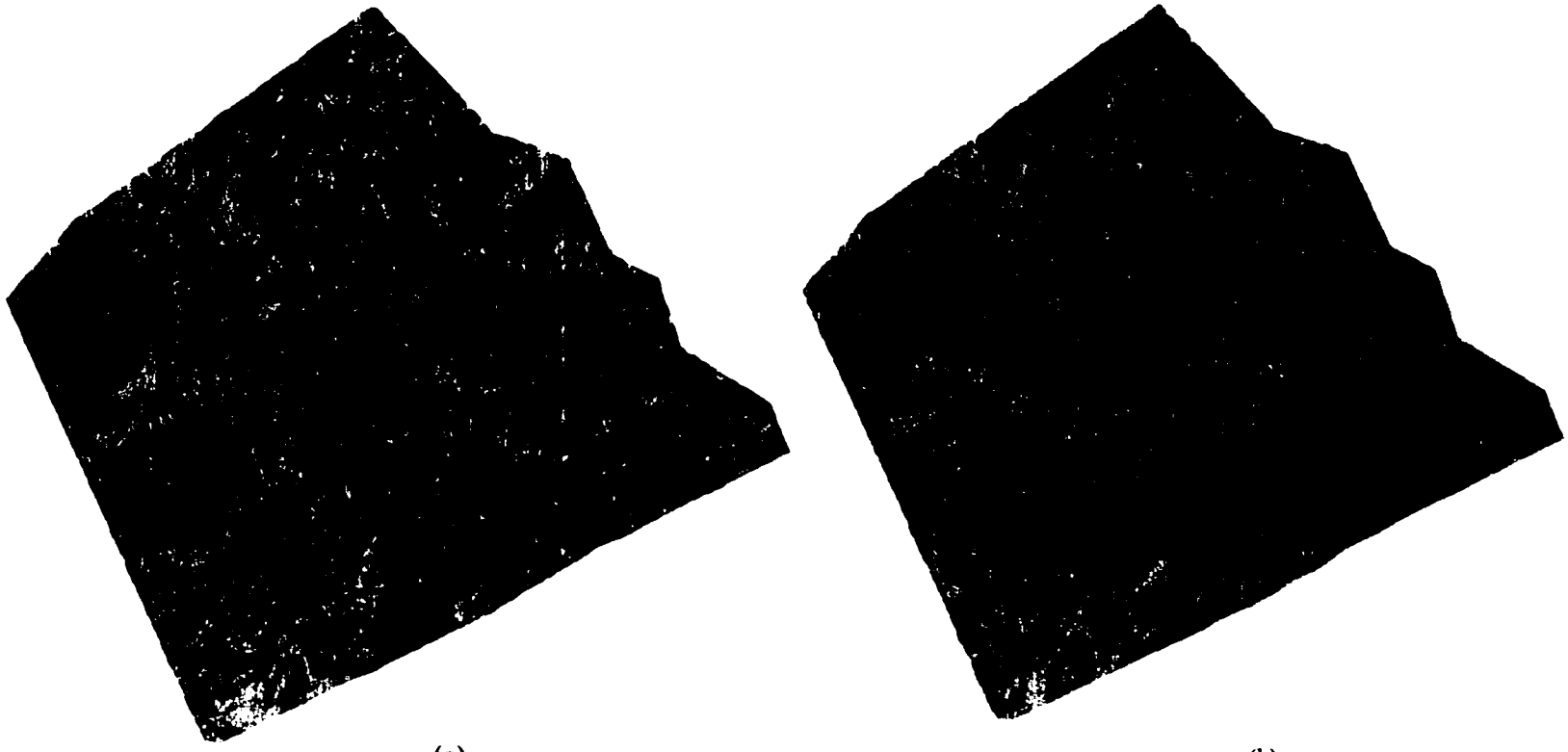


Fig. 7.9: Comparison of hill-shaded DEM derived from SAR stereo technique and topographic map. (a) shaded relief derived from SAR stereo technique; (b) shaded relief derived from 1:250,000 USGS topographic map.

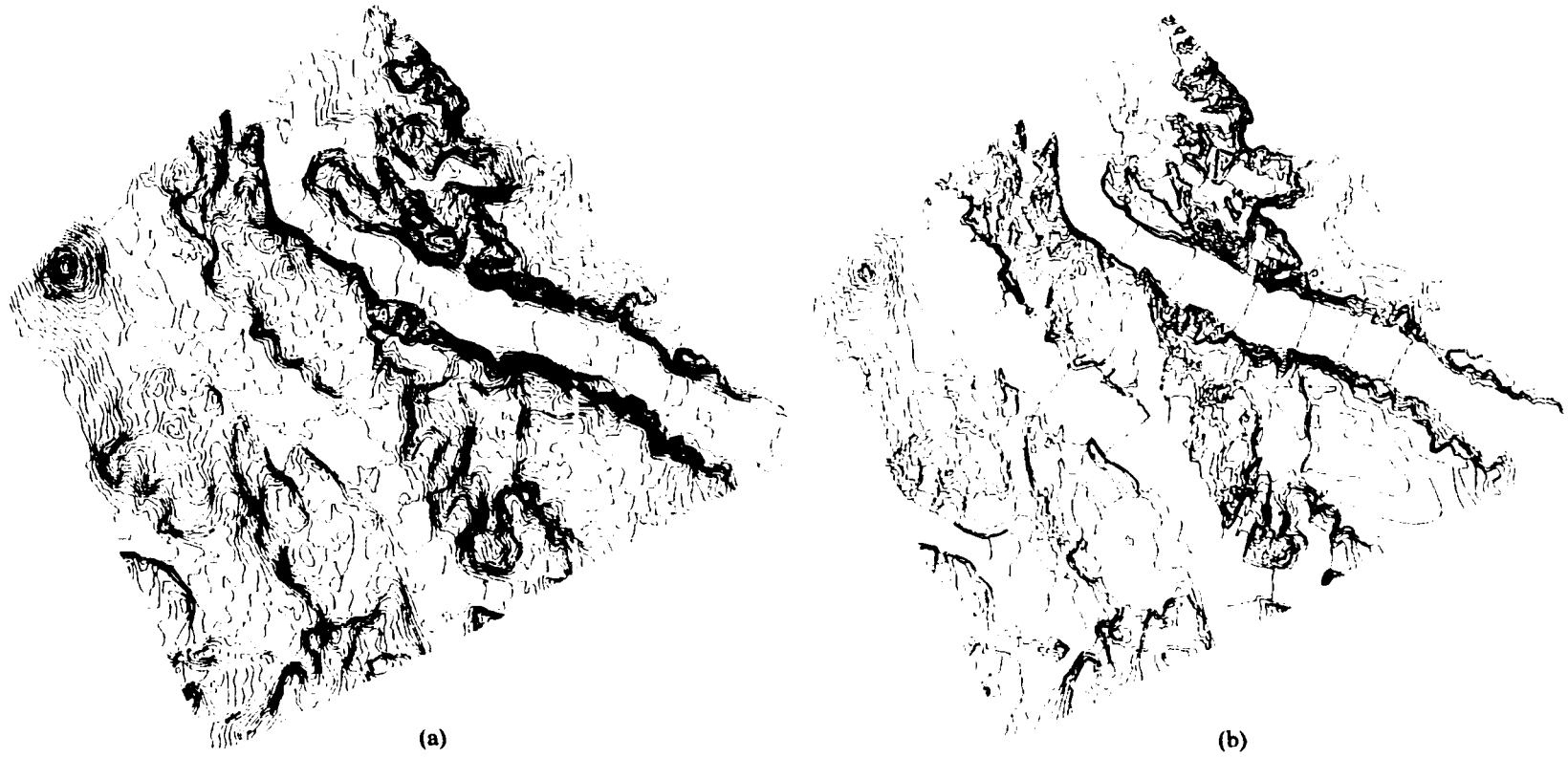


Fig. 7.10: Comparison of contours derived from SAR stereo with USGS 1:250,000 topographic map (a) contours from SAR stereo with an interval of 100 m; (b) USGS 1:250,000 topographic map with contour interval of 200 m.

from the 1:250,000 USGS topographic map and compared them with the SAR stereo derived values. The differences range from 2 to 154 m, and the RMSE is 77.5 m.

In summary, our test over the Terra Nova Bay area shows that the SAR stereo technique is capable of producing digital elevation data with spatial resolution superior to the 1:250,000 scale topographic map. The vertical accuracy of the SAR stereo derived elevation data can be about 70~80 m over the rugged mountainous regions like the Terra Nova Bay area of the Transantarctic Mountains. This experiment also reveals that the glacial floors and featureless ice plateaus and steep mountain slopes where layover and excessive foreshortening occur can fail the image matching process and cause significant data gaps. The data gaps in the flat ice floors and plateaus can be filled by data interpolation. To fill the gaps caused by the layover and excessive foreshortening, another stereo pair with a different looking direction should be processed.

7.4 Interferometric SAR technique

7.4.1 Theoretical background of SAR interferometry

The interferometric SAR (InSAR) technique has emerged as a precise approach to the extraction of high-resolution elevation data and measurement of small surface motion based on the phase information derived from the complex radar images (Zebker and Goldstein, 1986; Zebker, 1994). SAR interferometry combines complex radar signals (images) recorded by the antenna at slightly different locations to measure the phase differences between the complex radar images. There are three main ways to acquire SAR interferometric data: single-pass along-track, single-pass across-track, and repeat-pass across-track interferometry. At present, spaceborne SAR sensors, with only one

antenna, use the repeat-pass method to acquire interferometric data.

The repeat-pass interferometric method requires two or more passes of spacecraft in a precise repeat orbit such that the scene is imaged from almost the same, but slightly different, position and aspect angle during each pass (Zebker and Goldstein, 1986). If the ground is undisturbed between passes, the complex images will be highly correlated, and a large spatial baseline, which would not be feasible for a single platform, may be synthesized. The geometry of an interferometric SAR is shown in Fig. 7.11. An interferometer is formed by relating two complex SAR images of the same scene acquired at two orbital locations S_1 and S_2. As the radar signal transmitted by the SAR sensor is coherent, the complex SAR image possesses both phase and magnitude (quantities) information. For a distributed target a pixel (i,j) in two complex SAR images can be respectively represented as (Rodriguez and Martin, 1992; Ghiglia and Pritt, 1998):

$$s_1(i,j) = A_1(i,j)e^{j\phi_1} \tag{7.24}$$

$$\phi_1 = \frac{4\pi}{\lambda}\rho_1 \tag{7.25}$$

$$s_2(i,j) = A_2(i,j)e^{j\phi_2} \tag{7.26}$$

$$\phi_2 = \frac{4\pi}{\lambda}\rho_2 \tag{7.27}$$

where $s_1(i,j)$ and $s_2(i,j)$ are complex values of SAR image pixel at (i,j); A_1(i.j) and A_2(i,j) are the terrain reflectivity (backscattering) of two images; ρ_1 and ρ_2 are the ranges respectively from successive antenna S_1 and S_2 to the ground resolution element P(x,y,z) associated with image pixel (i,j); λ is the radar wavelength; and ϕ_1 and ϕ_2 are the phases of the returned radar signals in two images.

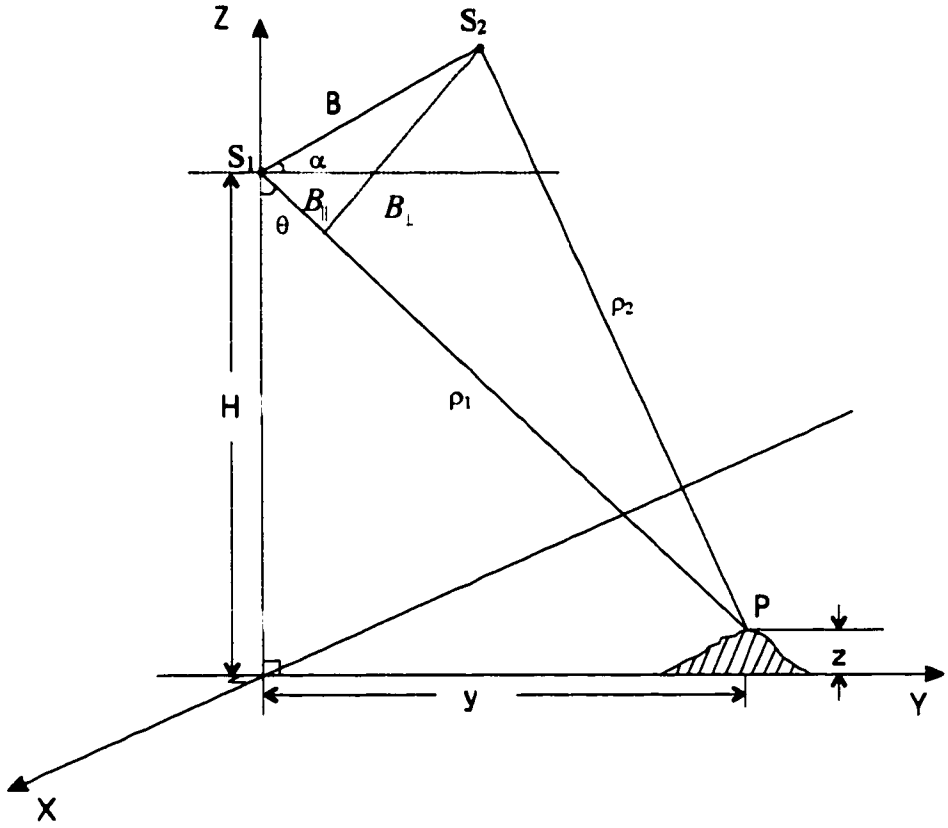


Fig. 7.11: Geometry of cross-track interferometric SAR. The SAR sensor is flying parallel to the z-axis(azimuth direction). The baseline B is determined by the sensor orbital positions S1 and S2 of two repeat passes. The baseline orientation is defined by the angle between baseline and horizontal line.

Then, the two complex images s_1 and s_2 can be interfered with each other by conjugate multiplication to form a complex interferogram. A pixel in the interferogram can be expressed as:

$$s_1(i,j)\cdot s_2^*(i,j) = A_1(i,j)A_2(i,j)e^{j\phi_M} \tag{7.28}$$

The constructive and destructive interference of coherent SAR images respectively recorded at orbital position S_1 and S_2 produce an interferogram with a two-dimensional fringe pattern. The measured phase in this interferogram is given by:

$$\phi_M = \text{mod}(\phi, 2\pi) \tag{7.29}$$

$$\phi = \phi_1 - \phi_2 = \frac{4\pi}{\lambda}\Delta\rho = \frac{4\pi}{\lambda}(\rho_1 - \rho_2) \tag{7.30}$$

After interfering the two complex SAR images, the phase ϕ_M measured from the interferogram is only the modulus of the absolute phase ϕ. To determine the range difference $\Delta\rho$, a phase unwrapping method must be used to convert the wrapped phase ϕ_M to the absolute phase. As the wrapped phase is measurable at each pixel to a fraction of a wavelength, interferometry is capable of high precision. But, it may suffer accuracy problems because of the modulo 2π ambiguity.

Having the range difference $\Delta\rho$ calculated by unwrapping the measured phase, along with the knowledge of the baseline length B and baseline angle α (Fig. 7.11), we can use the following equations to derive the surface elevation data (Zebker et al., 1994; Goldstein et al., 1989):

$$\sin(\theta - \alpha) = \frac{\rho_1^2 - \rho_2^2 + B^2}{2\rho_1 B} \approx \frac{\Delta\rho}{B} = \frac{\lambda\phi}{4\pi B} \tag{7.31}$$

$$\theta = \alpha + \arcsin(\frac{\Delta\rho}{B}) \tag{7.32}$$

$$y = \rho_1 \sin\theta \tag{7.33}$$

$$z = H - \rho_1 \cos\theta \tag{7.34}$$

where z is the elevation to be derived; B is the baseline, namely, the distance between two orbital position S_1 and S_2; α is the baseline angle with reference to the horizontal; θ is the look angle of the reference orbit S_1; ρ_1 and ρ_2 are the slant ranges of the reference orbital antenna S_1 and S_2 to the ground resolution element associated with image pixel (i,j); $\Delta\rho$ is the slant range difference between two orbit passes; H is the height of the reference orbital antenna S_1 above the datum; and ϕ is the unwrapped phase.

The accuracy of elevation estimated from interferometric SAR relies on baseline length and orientation measurement error due to the uncertainty of satellite orbital determination, phase measurement and unwrapping errors due to the thermal noise and geometrical distortion of complex SAR images, such as layover and shadows (Lin et al, 1994). The standard deviation of the elevation σ_z is related to the phase measurement error σ_ϕ by (Zebker et al., 1994; Madsen and Zebker, 1998):

$$\sigma_z = \frac{\lambda\rho_1 \sin\theta}{4\pi B_\perp}\sigma_\phi \tag{7.35}$$

$$B_\perp = B\cos(\theta - \alpha) \tag{7.36}$$

$$B_\parallel = B\sin\cos(\theta - \alpha) \tag{7.37}$$

where $B_\perp$, $B_\parallel$ are baseline components, respectively perpendicular or parallel to the range look direction (Fig. 7.11). The phase error can be approximated by (Madsen and Zebker, 1998):

$$\sigma_\phi = \sqrt{\frac{1}{SNR}} \tag{7.38}$$

where SNR is the signal-to-noise ratio. The repeat-pass interferometry is very sensitive to the baseline errors. The baseline geometry solely derived from satellite ephemeris data is not accurate enough and often causes unacceptable errors. Therefore, GCP points are needed to refine the baseline length and orientation. For optimal system performance, the baseline should be long enough so as to give sufficient phase sensitivity to elevation variation in terrain. The sensitivity of the interferometer to surface relief is dependent on the magnitude of the perpendicular component of the baseline $B_\perp$. The accuracy of extracted elevation data increases with a longer baseline. However, as the baseline becomes longer, the measured phase difference at two antennas will become larger than a wavelength and result in baseline decorrelation. The parallel component of baseline $B_\parallel$ does not affect the baseline decorrelation. The ERS-1 critical baseline is 1115 m, and the optimum baseline for elevation extraction would be about 200 m, and minimum baseline would be 75 m (Zebker et al., 1994).

A major limitation of repeat-pass interferometry is temporal decorrelation due to the variability of environmental factors. In addition, during the time span of revisit, surface motion may occur. If so, the measured phase is not only affected by the topography but also surface motion, and the topography-only phase can be expressed as:

$$\phi_{topo} = \frac{4\pi}{\lambda}(\rho_1 - \rho_2) - \phi_{baseline} - \frac{4\pi}{\lambda}\Delta m_r \qquad (7.39)$$

where Δm_r is the surface motion (displacement) along the slant-range look direction. When more than two passes of SAR data are available, differential interferometry can be used to separate the topography component and surface motion component by a double difference technique (Kwok and Fahnestock, 1996; Joughin et al, 1996).

7.4.2 Elevation data extraction with ERS Tandem data

A repeat-pass satellite interferometric technique was applied to a pair of ERS-1/ERS-2 tandem data over the Dry Valley region of the Antarctic (Fig. 7.12a). The tandem pair consists of two complex image frames-frame 660 with ERS-1 orbit 2387703 acquired on February 7, 1996, and frame 660 with ERS-2 orbit 4204 acquired on February 8, 1996, covering an area of 100 km by 100 km. The incidence angles of the ERS-1 and ERS-2 images at the scene center are respectively 23.23° and 23.22°. The perpendicular baseline component is about 190 m and parallel component is about 102 m. The one-day time span of the ERS tandem mission minimizes the temporal decorrelation and the surface motion contribution to the measured phase. Therefore, this would be suitable for extracting surface topography information. The terrain surface of the test area is characterized by rugged glacial landforms as seen in Fig. 7.12.

We used the ASF InSAR software (Guritz, 1998) to perform the interferometric SAR analysis on this tandem pair. Two raw CCSD (Computer Compatible Signal Data) files were first processed into SLC (Single Look Complex) images with the high precision phase preserving processing algorithm of the ASF software. Each of the SLC images has

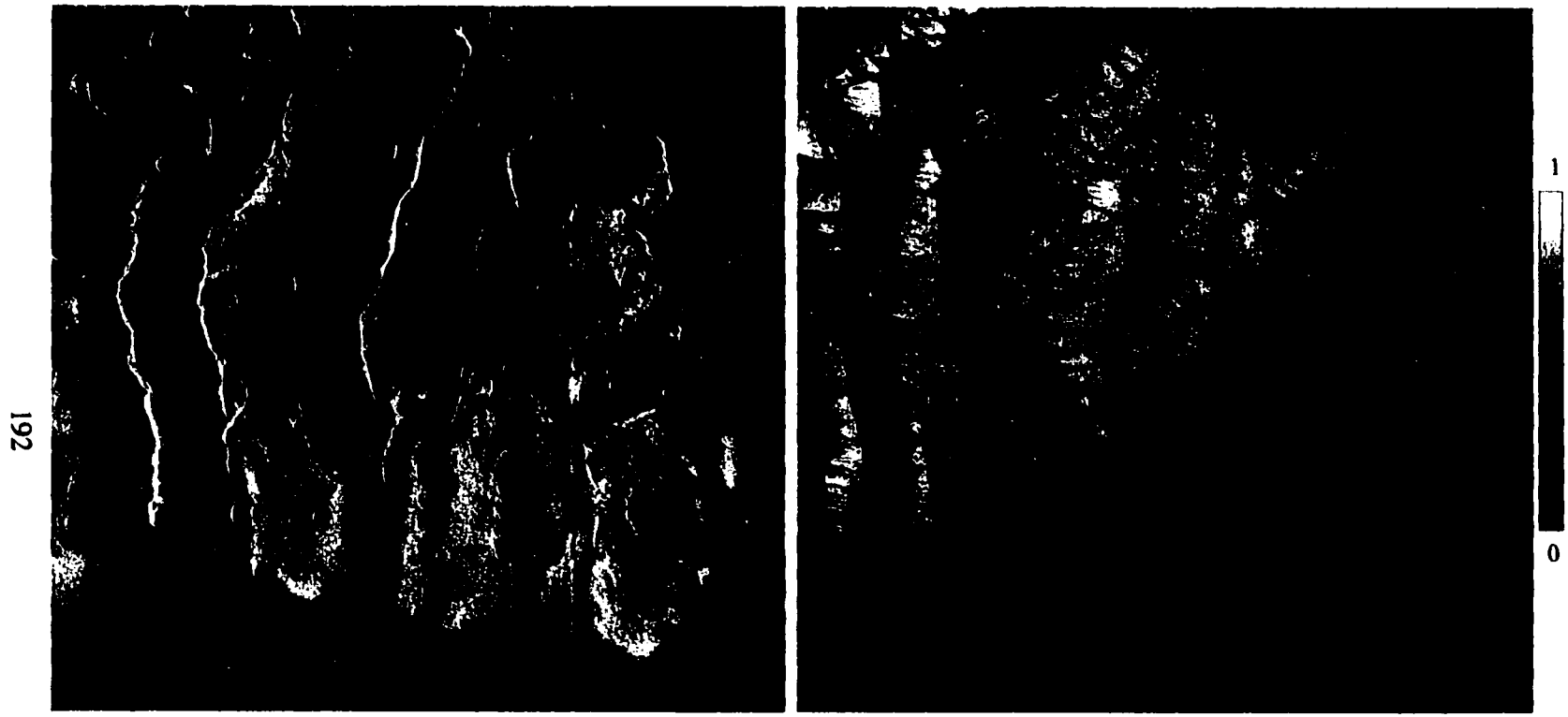


Fig. 7.12: Amplitude and correlation map of InSAR pair over Dry Valley, Antarctica. (a) Amplitude image of one single look complex image in the InSAR pair; (b) Correlation map of InSAR pair. Decorrelations are obvious in ocean, lakes and the layover area. The gray tones correspond to the calculated correlation coefficient, ranging from 0 to 1.

23415 lines and 4928 samples.

Subsequently, we conducted a series of interferometric processing steps on the two SLC complex images, including sub-pixel level co-registration, interferogram and coherence calculation, phase flattening, phase filtering, phase unwrapping, baseline refinement, seed point preparation, and unwrapped phase-to-elevation conversion.

For precise co-registration of the two SLC images, orbital meta data are first used to estimate the approximate offsets between the two complex images by a FFT (Fast Fourier Transform) matching. The gross offsets of the second image relative to the first one are estimated about 12 columns (range offset) and 125 rows (azimuth offset). Then, 20x20 regularly distributed points are selected. For each point, the phase coherence of several offsets near the gross offsets are computed between the two complex images, then a parabolic surface is fitted to search for the highest coherence offsets with sub-pixel level precision. Among 400 attempted points, 219 are successfully correlated with a fringe SNR above a specified threshold value. A bilinear plane is fitted on the 219 retained offset estimates. The second SLC image is resampled based on this fitted bilinear function in range and azimuth to co-register with the first SLC image with a 0.1 pixel accuracy.

With two co-registered complex images, the phase difference is determined on a pixel by pixel basis, and a single look interferogram is formed. To reduce phase noise, a multi-look interferogram is produced by box averaging 5 pixels in azimuth and 1 pixel in range. Under the assumption that the N complex looks have the same statistics, the maximum likelihood estimator of the phase is calculated from the single look complex values s_1 and s_2 (Massonnet and Rabaute, 1993; Small et al., 1995):

$$\phi_M = \arctan\left(\frac{\mathrm{Im}(\sum_{k=1}^{N} s_1^{(k)} s_2^{*(k)})}{\mathrm{Re}(\sum_{k=1}^{N} s_1^{(k)} s_2^{*(k)})}\right) \tag{7.40}$$

The box averaging with 5 azimuth looks results in a multi-look interferogram with 4683 lines and 4928 samples, and the phase noise was reduced by approximately $\sqrt{5}$ according to Rodriguez and Martin (1992). The wrapped phase contains the combined effects of baseline and topography.

An estimate of the level of coherence between the SLC images enables a determination of the expected usefulness and relative accuracy of this interferometric pair. The complex correlation coefficient was calculated based on the raw single look interferogram (Small et al., 1995):

$$\gamma = \frac{\sum_{k=1}^{N} s_1^{(k)} s_2^{*(k)}}{\sqrt{\sum_{k=1}^{N} s_1^{(k)} s_1^{*(k)} \sum_{k=1}^{N} s_2^{(k)} s_2^{*(k)}}} \tag{7.41}$$

The calculated correlation coefficient ranges from 0 to 1.0 and is shown in Fig. 7.12b. Overall, the tandem pair exhibits sufficient correlation over large ground areas. However, a number of significantly decorrelated areas exist. Decorrelation patterns of this tandem pair are closely related to terrain topography. The steep mountain slopes facing the radar looking direction show excessive foreshortening and layover as can be observed in Fig. 7.12a. In these geometrically distorted areas, the phase is either ambiguous or completely lost, and hence the correlation is very low. The water is an incoherent reflector. The ocean water of McMurdo Sound, Lake Fryxell, Lake Bonney, Lake Vida, Lake Vanda, and Upper Victoria Lake are completely lacking any coherence

between the interfered signals. In addition, the lower part (close to the ocean) of the Ferrar Glacier floor and the flat ice plateau in the upper right corner of the image (Fig. 7.12b) also show relatively low coherence, which may be due to a temporal change of freeze/thawing conditions or simply low signal strength. Low correlation means a high phase measurement noise, imposing a challenge to the subsequent phase unwrapping.

To make the phase unwrapping easier, the range and azimuth phase trends (ramps) due to the orbital baseline are removed at each pixel using knowledge of the spacecraft geometry. After removal of the flat terrain phase fringes, the remaining phase is known as a "flattened" or "deramped" interferogram, which is mainly due to topography effects. As shown in Fig. 7.13, the fringe pattern of the flattened interferogram is similar to a contour map of the line-of-sight displacement caused by the terrain relief. The interferogram shows quite sharp fringes, which is wrapped, namely, modulo 2π.

To further facilitate the phase unwrapping, a simulated phase image was created from a coarse resolution DEM derived from the combination of 1:50,000 and 1:250,000 USGS topographic maps. By removing this low-frequency topography phase from the flattened interferogram, the fringe density of the interferogram is greatly reduced. Incorporation of this coarse elevation data set simplifies the phase unwrapping problem.

After subtracting the coarse topography phase, a Goldstein phase filter (Goldstein et al., 1988) was applied to the residual interferogram, smoothing the phase noise while keeping the fringe lines sharp. The filter first performs a Fourier transform of a small piece of the phase image, then raises the Fourier transformed data to some power, and finally conducts an inverse Fourier transformation. The idea is that the spectrum of a phase image contains strong line components and white noise. By exponentiating the

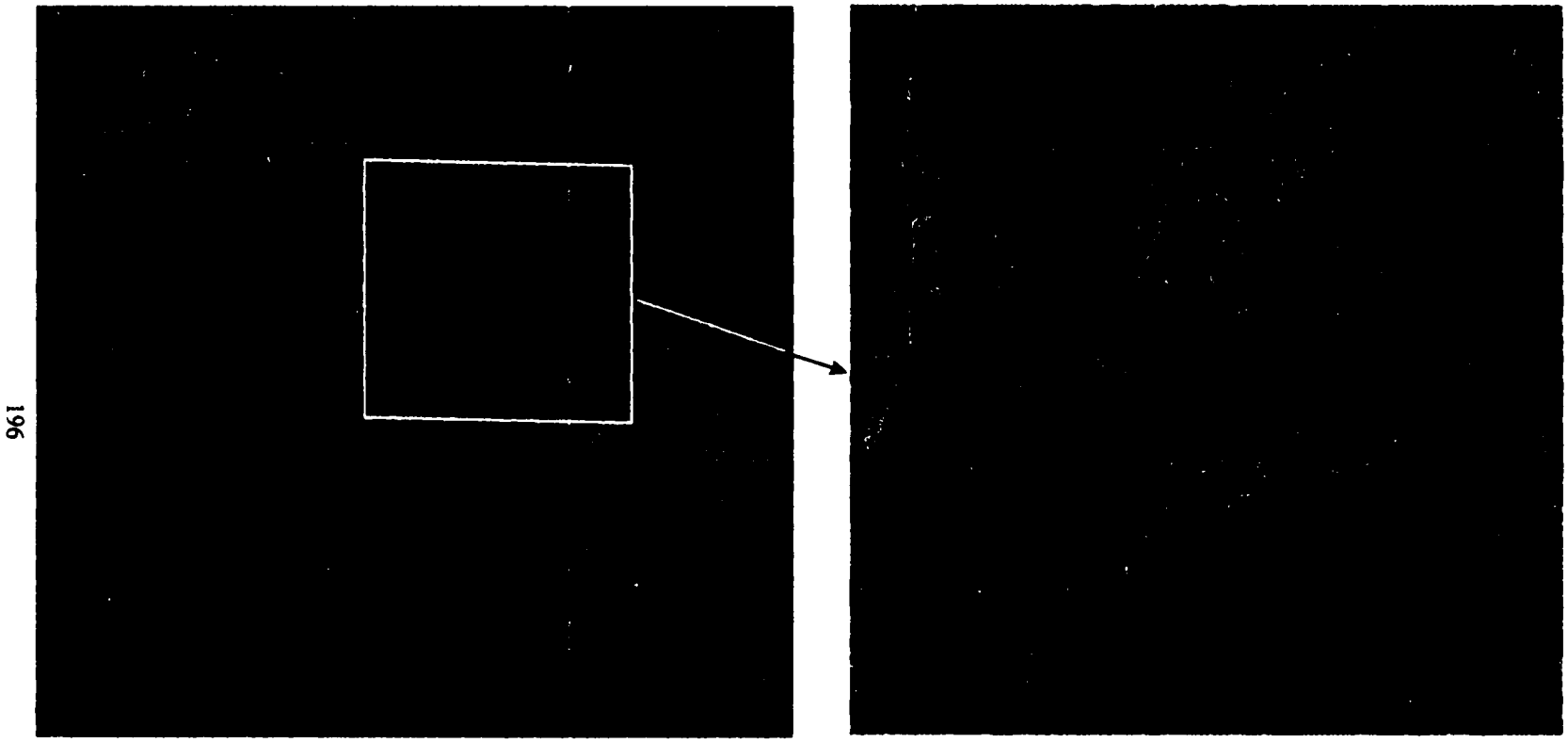


Fig. 7.13: Flattened (deramped) phase of interferogram

spectrum, the strong line components get stronger and the noise gets weaker, and therefore the signal in the phase image is preferentially amplified over noise.

Phase unwrapping is the critical step for interferometric SAR analysis. Goldstein's branch-cut algorithm (Goldstein et al., 1988) was used. As the measured phase is modulo of 2π, the correct multiple of 2π must be determined at each pixel to obtain an estimate of the actual phase with respect to an absolute datum. The resulting unwrapped phase is proportional to elevation. A fundamental assumption implicit in the phase unwrapping procedure is that the surface is relatively smooth and hence the unwrapped phase is also continuous, changing less than π radius from one adjacent point to the next. A phase unwrapping algorithm takes a starting seed location phase value, and adds or subtracts 2π as required to neighboring pixels, such that the phase difference does not exceed π from one pixel to the next. Goldstein's algorithm is an efficient and effective path following phase unwrapping procedure. The idea is to detect the inconsistencies (residues) in phase data and then connect nearby inconsistent points (residues) with branch cuts (walls) so that the residues are balanced and the sum of the cut lengths is minimized. Once the branch cuts are in place, the phase can be unwrapped along any path that does not cross the branch cuts.

Currently, the best operational orbit estimates for the ERS satellite are about 30 cm (Zebker et al, 1994). Despite this high precision, it is still not sufficiently accurate to estimate the baseline for InSAR analysis. Based on the orbital state vectors associated with each SLC image, the initial estimate of the baseline components are: $B_\perp = -192.325089\,\text{m}$, $\Delta B_\perp = -0.761062\,\text{m}$ (change over the entire

image), $B_{\parallel} = 102.473587$ m, and $\Delta B_{\parallel} = 0.137575$ m (change over the entire image). To refine the satellite orbital state vector and baseline geometry, four GCP points from 1:50,000 USGS map series and GPS measurements were used. First, the timing offset estimate is refined using the GCP points. The original information in the CCSD is inaccurate by 1.07 seconds in the along-track direction and 15.6 m in the across-track direction. To refine the baseline geometry, a series of tie points were generated from the coarse resolution DEM created from the USGS topographic map. To minimize the influence of surface motion, a mask was used to only pick up the tie points in the rock exposure area. The actual baseline varies slightly along track and cross-track. As these tie points have a known height, namely, the known unwrapped phase, the differences between the expected and observed phase at the tie point locations are used to improve the estimates of the baseline geometry, the orbit convergence phase rate, and the phase constant through an iterative non-linear least squares adjustment process. After refinement, the baseline component estimates are: $B_{\perp} = -189.826889$ m, $\Delta B_{\perp} = -4.865709$ m, $B_{\parallel} = 102.465057$ m, and $\Delta B_{\parallel} = 0.142298$ m. Incorporation of GCPs and tie points from a coarse resolution DEM improved the baseline parameters and hence reduced the absolute errors in the derived elevation data.

The low frequency topography phase was added back to the unwrapped phase. With the refined baseline parameters, the elevation was calculated at each integrated pixel in the scene, except for those marked during the coherence calculation and phase unwrapping steps as lacking reliable phase information. The tie point file used for the baseline refinement was used here as the seed points for converting unwrapped phase to elevation.

InSAR derived elevation data are shown as a hill-shaded relief image in Fig. 7.14. A number of significant data holes exist, which mainly correspond to lakes, ocean water and the steep mountain slopes in our test scene. In these areas, phase coherence is lost (Fig. 7.12b), and phase discontinuity actually occurs. The high density of residues and corresponding branch cuts isolated these areas from the rest of the interferogram, making phase unwrapping unreachable and impossible. In the slant range projection, 72.1% pixels are successfully unwrapped and integrated. In other words, data gaps account for 27.9% pixels of the scene, including ocean and lake water. In comparison with the DEM derived from the combination of 1:50,000 and 1:250,000 topographic maps (Fig. 7.15), it is quite clear that the shapes of terrain features present in the InSAR derived elevation data are in good agreement with those derived from the USGS topographic maps, and that the InSAR derived elevation data contain more topographic details than the 1:50,000 USGS topographic maps, let alone the 1:250,000 maps.

To evaluate the absolute vertical accuracy, the elevation values derived from the InSAR technique were compared with those derived from 1:50,000 USGS topographic maps along four cross-sections (Fig. 7.16). The RMSE values for these cross-sections are respectively 19.4 m, 20.8 m, 26.2 m, and 36 m. These RMSE values approximately reflect the level of the vertical accuracy that the InSAR technique can achieve in the rugged glacial mountains, but they do not represent a precise accuracy assessment of the InSAR technique because the accuracy of 1:50,000 topographic map itself is limited. In some areas, however, systematic errors of about ±50 m, ±100 m were detected in the InSAR derived elevation data. The systematic errors were caused by the integer ambiguity of the phase unwrapping. Using equation (7.35), the elevation difference of

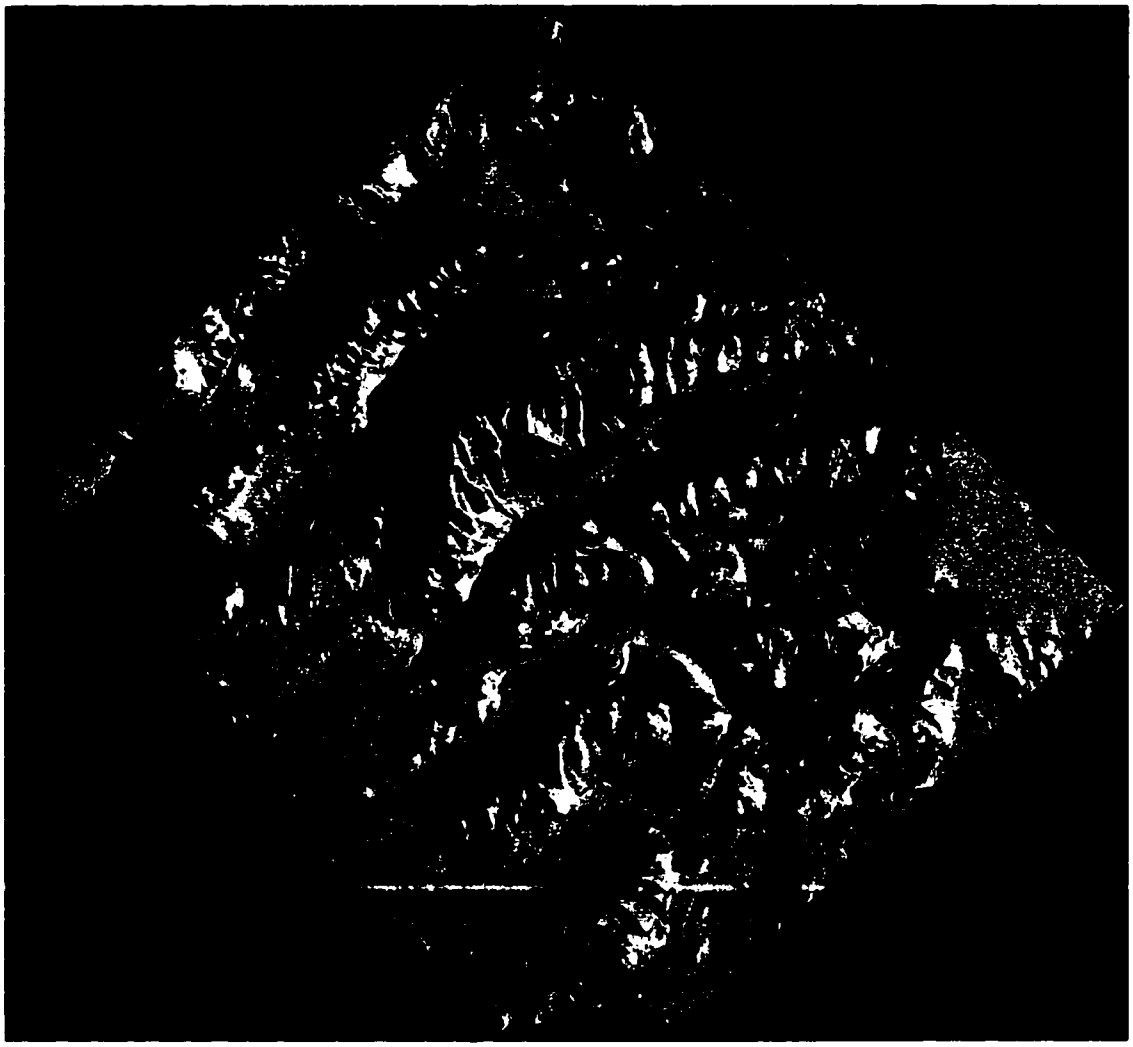


Fig. 7.14: Hill-shaded DEM for the Dry Valley derived from

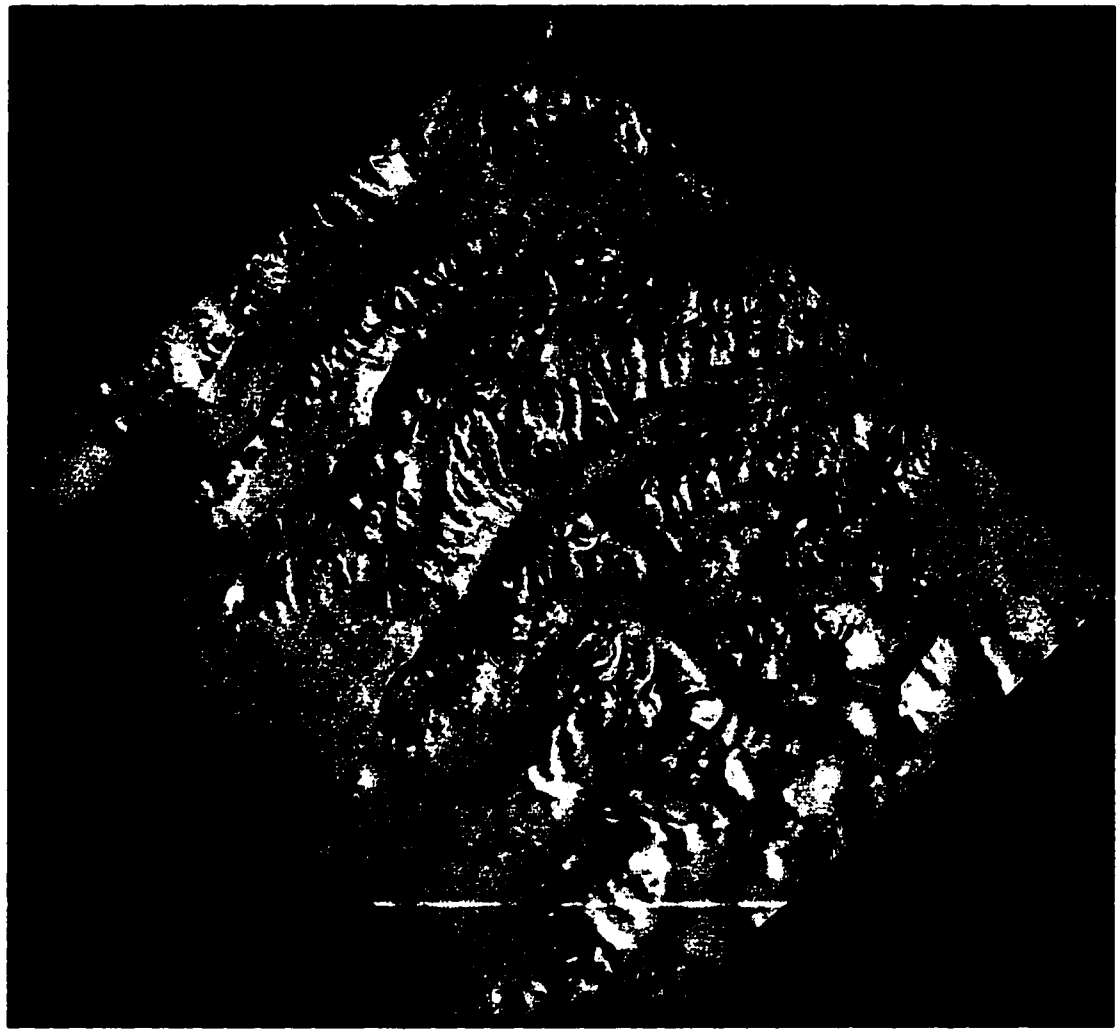
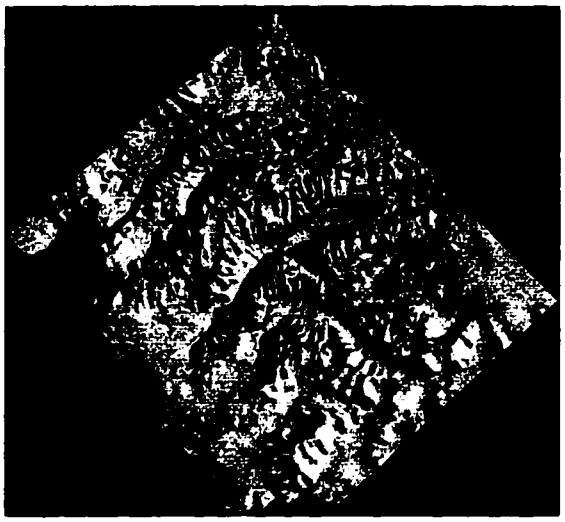


Fig. 7.15: Hill-shaded DEM derived from the combination of 1:50,000 and 1:250,000 USGS topographic maps



Location of cross sections

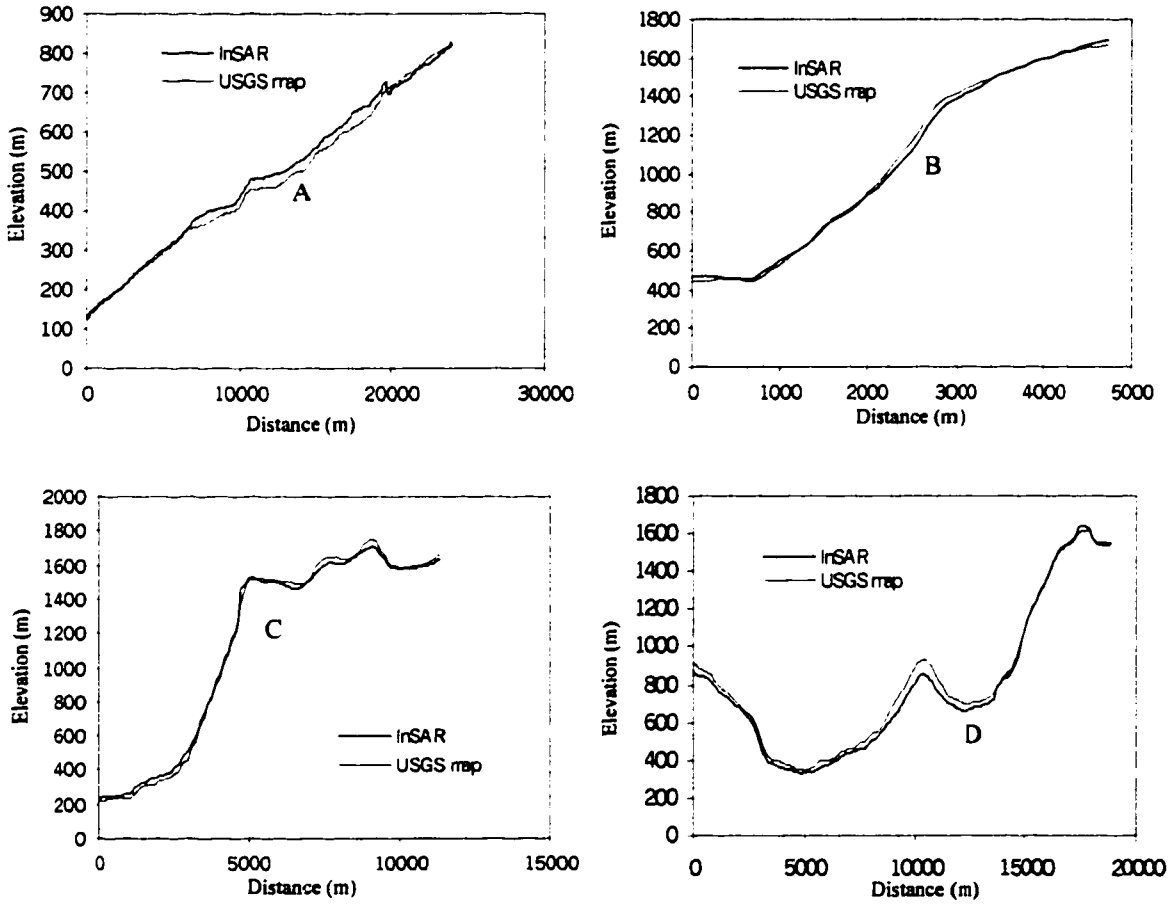


Fig. 7.16: Accuracy assessment of InSAR derived elevation data

one interferometric fringe (2π) was calculated for our InSAR pair. It is 50.4 m. Then, the systematic errors were corrected in a GIS environment by adding or subtracting 50.4 m or 100.8 m in these areas.

In summary, our experiment in the Dry Valley area shows that the spaceborne interferometric SAR technique is capable of extracting elevation data with a higher spatial resolution than 1:50,000 USGS topographic maps. A vertical accuracy of about 20~30 m can be achieved. It is also shown that it is impossible to get a full coverage of elevation data with one InSAR pair in the rugged mountainous area. Due to the low incidence angle of ERS tandem data (23°) and the high surface slopes of the Dry Valley, a considerable portion of the scene is affected by excessive foreshortening and layover, causing the data holes. To fill the holes, at least another InSAR pair with an opposite-looking direction (descending in our case) is required. The other observation from our experiment is that the automated phase unwrapping procedure tends to cause systematic errors in mountainous regions. To avoid the systematic errors and enlarge the phase unwrapping area, a GIS-based interactive phase unwrapping method is required.

7.5 Technical strategy and terrain characteristics

The performances of different spaceborne elevation acquisition techniques are influenced by numerous factors, including orbital geometry, sensor properties, environmental conditions and terrain complexity. Radar and laser altimeters are nadir-pointing active sensors designed to measure surface height by tracking the shape and timing of the returned radar or laser signal. Both radar and laser altimeters are capable of operating day and night, generating traverses of point elevation measurements. Although the

performance of laser altimeter data may deteriorate in the condition of bad weather, such as clouds and precipitation, it generally provides much more accurate measurements than the radar altimeter due to the small footprint of the laser beam. Stereo and shape-from-shading technique can be applied to both optical and radar image data. Optical sensors have a near-vertical imaging geometry and moving perspective projection, therefore, the geometric integrity and fidelity of optical images is much better than the images acquired by SAR sensors that use side-looking geometry and slant range projection. Optical images often suffer from clouds, while SAR images are free of the influence of atmospheric conditions. Nevertheless, SAR images have much more noise and speckle than optical images. Geometric distortion, clouds, and image noise would adversely influence stereo and shape-from-shading techniques. In addition, the application of shape-from-shading to SAR images over the snow-covered regions might be complicated by the influence of snow grain size (Sohn, et al., 1999). The interferometry technique requires the coherent signal and is only applicable to complex SAR images. The quality of elevation data derived from spaceborne interferometric SAR technique is affected by the phase noise, geometric distortion, baseline error and temporal decorrelation.

Among the influencing factors, sensor properties (spatial resolution, signal-to-noise ratio) have been improving with the advance of technology. The imaging geometry can be selected or adjusted through the mission plan. Microwave sensors can overcome the adverse atmospheric effects to a large extent. However, the influence of terrain condition is relatively persistent and complex. Therefore, it is important to understand the relationship between the performance of various spaceborne techniques and terrain characteristics, namely, the geography of the spaceborne techniques.

From our experimental results in the preceding sections and the review of the literature, it is quite clear that the slope, curvature, tonal and textural properties of terrain surface affect and limit the accuracy and density of elevation data that can be achieved by spaceborne techniques. By using a prior knowledge about the terrain of Antarctica, we can design and select an appropriate technique or a possible combination of several techniques for improving the spatial resolution and vertical accuracy of our DEM. Based on the analysis of the surface morphology of the Antarctic (Fig. 6.1) and visual inspection of tonal and textural properties of the terrain from AVHRR images (Ferrigno et al., 1996) and Radarsat SAR images (Jezek, 1998), the suitability and feasibility of each spaceborne technique are evaluated in terms of terrain conditions of the Antarctic (Table 7.1).

Guided by our evaluations, we formulated a technical strategy for enhancing our Antarctic DEM in the future. The periphery ice shelves, including Ross Ice Shelf, Amery Ice Shelf, Ronne-Filchner Ice Shelf and other small ice shelves, are flat or slightly tilted. The GLAS laser altimeter will be the most appropriate instrument to acquire the accurate and high-resolution elevation data in the near future. The ice rises, ice-covered islands, ice domes and interior ice sheet are characterized by low surface slopes, and slight or moderate undulations. There are tonal variations but a lack of features and texture information in both optical and SAR images. For these areas, neither optical nor SAR stereo will work well. Laser altimeter and shape-from-shading techniques are well-suited for this type of terrain. Laser altimetry produces accurate absolute height measurements along satellite tracks, whereas the shape-from-shading technique produces dense two-dimensional relative measurements of local height differences that require boundary values to produce actual heights for a terrain model. The combination of laser altimetry

Spaceborne Technique	Favorable Terrain Conditions	Adverse Terrain & Weather Conditions	Applicable Areas in Antarctica	Vertical Accuracy	Spatial Resolution
Radar Altimeter	Low surface slopes < 1°	Rugged terrain	Ice shelves and interior ice sheet	2 m for ice shelves; 15~30 m for ice sheet	335 m along track; about 5 km between tracks
Laser Altimeter	Low or moderate surface slopes < 30°	High-relief, clouds, rain, or snow conditions	Ice shelves, interior ice sheet, smooth land	15 cm for < 1° 22 cm for < 3° 1 m for < 10° 10 m for < 30° (Schutz, 1998)	170 m along track; about 1 km between tracks
Optical Shape from Shading	Uniform snow or ice cover, smooth terrain surfaces	Solar shadows, nunataks, dust-covered snow, hoar frost, wet snow, ice crevasses	Ice rises, ice domes, ice and snow-covered islands, undulating interior ice sheet	2~4 m	Image resolution cell level
SAR Shape-from-shading	Uniform snow or ice cover, smooth and undulating surface	Radar shadows, radar speckles, nunataks, wet snow, ice crevasses; snow grain size variation	Ice rises, ice domes, ice and snow-covered islands, undulating interior ice sheet	20 m (Thomas et al., 1991)	Image resolution cell level
Optical Stereo	Abundant non-periodic terrain features and textures	Featureless, low or repetitive textural pattern, surface breaklines (abrupt slope change), clouds	Hilly coastal regions; Rugged mountainous regions (with multiple stereo pairs acquired at different times).	6~18 m for favorable base-to-height ratio (Vincent et al., 1987; Ehlers et al., 1989)	Depending on the density of successfully matched points
SAR Stereo	Abundant terrain features and textures, low to moderate terrain relief	Featureless, poor texture, excessive foreshortening and layover, radar shadow	Hilly coastal regions; Rugged mountainous regions with descending and ascending stereo pairs	10~50m for flat or moderately rugged terrain (Marra etal., 1998); 70~80 m for rugged terrain	Depending on the density of successfully matched points
InSAR	Low to moderate terrain relief, minimal surface motion	Excessive foreshortening, layover, radar shadows	Hilly coastal regions; Rugged regions with descending and ascending InSAR pairs	5~20 m for flat and moderately rugged terrain (Marra et al., 1998); 20~30 m for rugged terrain	Image resolution cell level

Table 7.1: Suitability of different spaceborne techniques in terms of terrain types

and shape-from-shading techniques promises a terrain model that has higher horizontal resolution and better vertical accuracy than either technique could create alone.

Coastal regions, including Prince Charles Mountains in Lambertian Glacial basin, Sør Rondane Mountains, Sverdrup Mountains and other mountains in Queen Maud land, Pensacola Mountains, Shackleton Range, the volcanoes in the West Antarctica, and Palmer Land of Antarctic Peninsula, have a moderate relief but are not too rugged. Terrain features and textural information are abundant in satellite images. Despite the presence of terrain-induced geometric distortions, the distortions are not very serious. Therefore, optical and SAR stereo, and interferometric SAR technique can be effectively applied to these hilly and moderately rugged coastal regions to extract high-resolution elevation data.

The Transantarctic Mountains, Ellsworth Mountains, Alexander Island and Graham Land of the Antarctic Peninsula have steep and rugged terrain. High surface slopes and abrupt slope variations (surface curvature) have induced serious layover, excessive foreshortening and radar shadows in the SAR image, and radial terrain displacement and solar shadows in optical images. No single technique works very well for this type of high-relief terrain. SAR stereo technique uses the same-side geometric configuration. The image with a steep-looking angle in the stereo pair contains excessive foreshortening and layover areas, while the image with a shallow-looking angle in the stereo pair tends to have radar shadow areas. Poor or failed image matching often results in deteriorated quality or data holes. Similarly, an InSAR pair with a steep-looking angle suffers from the excessive foreshortening and layover as demonstrated, whereas the InSAR pair with a shallow-looking angle tends to have radar shadows, causing the decorrelation. In other

words, one stereo or InSAR pair can only derive elevation data for one side of mountain slopes and flat valley floors. The combination of multiple pairs with different looking directions (descending and ascending pairs) has the potential to obtain the full coverage of elevation data. Optical stereo has a relatively small geometric distortion and works better than SAR stereo and InSAR techniques in the high relief areas. Optical stereo techniques will be influenced by terrain induced shadows due to the low solar elevation angle in polar regions. In addition, rugged mountains of the Antarctic are frequently cloud-covered, and optical images tend to be affected by clouds. Therefore, to achieve a complete coverage, multiple optical stereo pairs acquired at different times or seasons might also be needed.

In the hilly coastal regions and rugged mountains, optical stereo, SAR stereo and InSAR techniques can be separately applied. Then, elevation data independently extracted from different techniques can be combined together during the interpolation process to create a complete terrain model. It should be noted that an existing low-resolution DEM, optical stereo, SAR stereo, and InSAR techniques can be also combined at the elevation extraction stage to eliminate some inherent ambiguities for an individual technique, hence increasing the overall accuracy and resolution of the resulting elevation model. For example, a coarse-resolution DEM could be used to predict and seed initial match offsets for the stereo technique so as to reduce the search space and the probability of finding a false match. As demonstrated in our InSAR experiment, a low-resolution DEM can be also used to refine the baseline and assist in the phase unwrapping process. Similarly, the stereo derived height measurements can help in the interferometric phase unwrapping by acting as seed points or by reducing the density of interferometric fringes.

CHAPTER 8

CONCLUSIONS

This research addressed a spectrum of technical and theoretical issues encountered in the construction of a large-scale DEM with multiple data sources. Enormous data volume, various data sources and different terrain types complicate the large-scale DEM generation process and pose significant technical difficulties and challenges. To tackle them, we developed several new methods and also improved some existing algorithms. By using these newly developed and extended methods and algorithms along with available GIS functions, we integrated a variety of topographical data and constructed a complete, consistent, and high-resolution DEM over the Antarctic.

Based on internal analytical and graphical functions of the proprietary GIS software ARC/INFO, a software package, IGIAS, was developed to support and automate the large-scale DEM generation process. Through a graphical user interface, this software provides an easy access to both a topographical database and computational functions. Vector data can be directly incorporated into DEM processing, and the resulting raster-based DEM grids can be readily visualized and evaluated with reference to other spatial data layers. The available GIS functions in spatial data manipulation, map algebra, statistical analysis, mathematical modeling, spatial data display and visualization are very

useful for handling a variety of topographical data and facilitate the DEM computation over a large area. In the case of the Antarctic, we demonstrated that data import and conversion functions in ARC/INFO can be used to transform various types of topographical data into a coherent and topologically structured format that can be efficiently stored, retrieved and manipulated by the subsequent GIS operations. The map projection functions in the GIS software can be used to project different topographical data layers into the same planimetric coordinate system, while the manipulation of attribute tables of topographical data layers can adjust the height measurements using a common vertical reference datum. In addition, graphical and logical data selection functions, map algebra, vector data clipping and merging functions can be utilized to manage the topographical database and prepare the input data for interpolation.

Currently available GIS software does not provide an entire solution to the development and refinement of a large-scale DEM. Some available algorithms and methods do not provide satisfactory results, and some required functions are not available. We modified and extended the unsatisfactory conventional algorithms, and also developed several new external functions by using C language programming. To extract the elevation data from raw remotely sensed data, public domain software-ASF InSAR tool and commercial software-Vexcel SAR stereo toolkit were used.

A consistent and error-free topographical database is the prerequisite for constructing a high-quality large-scale DEM. GIS software does not have the functions for detecting and correcting errors in spatial data. To achieve a reliable and error-free topographical database, we proposed a comprehensive error detection scheme. It consists of three complementary methods. Global statistical analysis method identifies the global extreme

values by comparing each individual data value with the global statistics of the entire data set, and a cross-validation method detects systematic errors or random errors by comparing two overlapping data sets. Scientific visualization and image simulation methods can effectively detect local errors in topographical data by rendering and inspecting topographical surfaces in various ways, especially through analytical hill-shading and a synthetic stereo view. It is demonstrated that this method is also useful for detecting the positional errors of topographic data by comparing simulated images with real satellite images. A new locally adaptive and robust statistical analysis method is developed in this research to detect the subtle local errors in an irregularly distributed spatial data set based on the surface continuity and smoothness assumptions. These error detection methods were applied to various input source data and demonstrated their effectiveness and sharpness. Due to these extensive and rigorous error checking procedures, we believe that our Antarctic DEM is highly reliable and free of gross errors and blunders relative to the source data.

Spatial interpolation represents the primary computation involved in the large-scale DEM generation. We compared different interpolation algorithms in terms of type and distributional pattern of input data. The currently available algorithms in ARC/INFO only give reasonable interpolation results for evenly distributed point measurements. For traverse data and contour-based data, no single interpolation algorithm can provide a satisfactory result. The failure of available algorithms on traverse data is mainly due to the fact that they cannot effectively handle an anisotropic distribution of neighboring points and very dense data along flight tracks. We designed a new approach to the interpolation of traverse data. It consists of reducing and filtering the original

measurements along tracks by using a super-block based searching process and a median filter, interpolating the reduced data into a coarse grid using a quadrant neighborhood based Inverse Distance Weighted algorithm, and then interpolating the TIN model, created by combining the reduced data and the coarse grid points, into a fine DEM grid. It is demonstrated that this approach gives a more reasonable result for interpolating traverse radar data. For the interpolation of contour-based topographical data, we designed a two-stage TOPOGRID algorithm. This algorithm can incorporate contours, surface structural lines and spot elevation points during interpolation, and preserves the topographical details present in the source data, while avoiding the introduction of artifacts. In the existing literature, grid spacing for interpolation is often subjectively or arbitrarily determined. In this research, FFT-based power spectrum analysis and multiple resolution analysis are introduced to determine an appropriate interpolation grid spacing by diagnosing the topographic information content in the source data.

Data integration and merging operations are required for the creation of a large-scale DEM. This research divided the topographical data into competitive data and complementary data. To exploit the best elements of the available topographical data, a set of rules (criteria) were developed in the case of the Antarctic topographical database to select the competitive topographic source data based on the quality assessment. It is also demonstrated that spatially and functionally complementary topographical data sources can be synergistically combined during the interpolation stage to create a more reliable and accurate elevation model than what a single source could create alone. Using a Hermite blending function, we developed a method to seamlessly merge individual DEM blocks derived from different data sources along an irregular buffer zone. This

data merging method successfully improved the consistency and transition between different data sources. Visual inspection of the shaded relief image of the DEM and the ice drainage pattern extracted from the DEM show that the Antarctic DEM is seamless and geomorphologically consistent throughout the continent. No abrupt changes or "data cliffs" exist along the seam lines between different data sources.

As the case study, we created a high resolution Antarctic DEM by integrating a variety of topographical data in a GIS environment, including satellite radar altimeter data, the most detailed cartographic data digitized from topographical maps, and airborne radar data. The overall quality of the resultant DEM is evaluated in terms of the horizontal resolution, positional precision, vertical accuracy, reliability and topologic consistency. The DEM has a variable horizontal resolution, approximately 200 m over the Antarctic Peninsula and the Transantarctic Mountains, 400 m in the coastal regions, and approximately 5 km in the interior. It represents a substantial improvement over previous efforts, especially in mountainous and coastal areas. Comparisons with high-resolution satellite images at different scales demonstrate that our DEM not only describes the broad topographic configuration and the true ice drainage pattern, but also captures an unprecedented detail of the terrain surface. With reference to large scale topographic maps and several traverses of GPS, geodetic leveling and airborne radar data, the vertical accuracy of the DEM is estimated at about 100-130 m over rugged mountainous areas, better than 2 m for the ice shelves and better than 15 m for the gently sloping interior ice sheet. Accuracy is estimated at about 35 m for the relatively rough and steeply sloped portions of the ice sheet perimeter. The Fourier analysis and directional variogram analysis show that the errors of the contour-derived DEM data have a strong spatial

autocorrelation. The autocorrelation strength and range are scale-dependent and anisotropic, and are closely related to the shape, orientation and wavelength of the topographical features.

Of all the world's continents, Antarctica is the coldest, the highest, and the least known. Conventional maps at a scale of 1:250,000 or larger only cover about 20% of the area of Antarctica (Swithinbank, 1988). In our DEM, the Prince Charles Mountains in the Lambert Glacier Basin and the flat ice plateau inside the 81.4°S circle are two known areas of poor quality, where the best existing data source is small scale contour data of dubious accuracy. The present Antarctic elevation model can be improved by introducing more reliable and accurate topographical source data. Given the large ground coverage and relatively low-cost, advanced spaceborne remote sensing techniques promise an effective and efficient approach to the improvement of available large-scale DEMs. In the case of the Antarctic, we point out that the GLAS Laser Altimeter in the near future will provide much more accurate and high resolution measurements over the ice shelves and interior ice sheet with low and moderate surface slopes. The experiment with a panchromatic SPOT image over the Crary Ice Rise shows that a shape-from-shading technique is useful for enhancing the spatial resolution of the elevation data over ice rises, ice covered islands, ice domes and an undulating ice sheet. The initial experiments in the Terra Nova Bay area and the Dry Valley region show that SAR stereo and InSAR techniques have great potential for further enhancing the resolution and accuracy of our DEM in the hilly coastal regions and rugged mountainous regions.

From this research, we conclude that GIS technology provides an effective means of integrating not only diverse topographical data sources, but also various alternative

computational methods and techniques for creating a complete, consistent, and high resolution DEM over a large area.

The initial contribution of this research to the methodological literature include the error detection and correction methods, traverse and contour data interpolation procedures, determination of optimal interpolation intervals, a data merging scheme, error pattern analysis methods, ice flow line extraction algorithm, and experimental results with shape-from-shading, SAR stereo and InSAR techniques. These newly developed and extended methods and techniques, along with available GIS functions, collectively constitute a comprehensive means for handling the technical issues raised in a large-scale DEM generation process.

The major substantive contribution of this research is a new high-resolution Antarctic DEM. We believe that our Antarctic DEM product represents the most detailed and accurate digital description of the Antarctic topography published to date. It will be a very useful data source for mapping and studying the geomorphologic characteristics and dynamic behaviors of the Antarctic. The immediate use of this DEM will be terrain distortion corrections of Radarsat SAR imagery for producing the first ever, complete ortho-rectified image mosaic of the Antarctic (Jezek et al., 1998). We have established a GIS-based database to maintain all the topographic source data for each portion of the DEM. Consequently, as new topographic data become available, it is practical to continually update the DEM and guarantee that our DEM product is the best and most timely representation of the Antarctica topography.

LIST OF REFERENCES

Ackermann, F. (1996). "Techniques and strategies for DEM generation," in C. Greve (ed) *Digital Photogrammetry: An Addendum to the Manual of Photogrammetry*, American Society for Photogrammetry & Remote Sensing.

Akima, H. (1978). "A method for bivariate interpolation and smooth surface fitting for irregularly distributed data points," *ACM Trans. On Mathematical Software*, 4 (2), 148-159.

Al-Daoud, M., and S. Roberts (1996). "Applying efficient techniques for finding nearest neighbors in GIS applications," In D. Parker, Taylor & Francis (eds) *Innovations in GIS 3*, pp. 95-103.

Ansoult, M. M. (1989). "Circular Sampling for Fourier Analysis of Digital Terrain Data." *Math. Geol.* 21 (4), 401-410.

Auerbach, S. and H. Schaeben (1990). "Surface Representations Reproducing Given Digitized Contour Lines," *Math. Geol.*, 22 (6), 723-742.

Aumann, G., H. Ebner and L. Tang (1991). "Automatic derivation of skeleton lines from digitized contours," ISPRS Journal of Photogrammetry and Remote Sensing, 46, p259-268.

Bailey, T. C. (1994). "A review of statistical analysis in geographical information systems," in Fotheringham, S., and Rogerson, P., (eds) *Spatial analysis and GIS*, Taylor & Francis.

Bailey, T. C. and A. C. Gatrell (1995). *Interactive Spatial Data Analysis*, Longman Scientific & Technical, 413 p.

Balce, A. E. (1987). "Determination of optimum sampling interval in grid digital elevation models (DEM) data acquisition." *Photogrammetric Engineering & Remote Sensing* 53(3): 323-330.

Bamber, J. (1994). "A digital elevation model of the Antarctic ice sheet derived from ERS-1 altimeter data and comparison with terrestrial measurements," *Ann. Glaciol*, 20, 48-54.

Bamber, J. and C. R. Bentley (1994). "A comparison of satellite-altimetry and ice-thickness measurements of the Ross Ice Shelf, Antarctica," *Ann. Glaciol*, 20, 357-364.

Bamber, J. L. and R. A. Bindschadler (1997). "An improved elevation data set for climate and ice-sheet modeling: validation with satellite imagery," *Ann. Glaciol*, 25.

Barnet, V., and T. Lewis (1984). *Outliers in Statistical Data* (Second Edition, John Wiley & Sons).

BAS, SPRI, and WCMC (1993). *Antarctic Digital Database User's Guide and Reference Manual*. Scientific Committee on Antarctic Research, Cambridge.

Batson, R.M., K. Edwards, E.M. Eliason (1976). "Synthetic stereo and LANDSAT pictures," *Photogrammetric Engineering & Remote Sensing*, Vol. 42, No. 10, pp1279-1284.

Bentley, C. R. and K. C. Jezek (1981). "RISS, RISP and RIGGS: Post-IGY glaciological investigations of the Ross Ice Shelf in the U.S. programme," *Journal of the Royal Society of New Zealand*, 11, 355-372.

Bentley, J.L., B.W. Weide, and A.C. Yao (1980). "Optimal expected-time algorithms for closest point problems," *ACM Trans. On Mathematical Software*, 6, 563-580.

Bindschadler, R. and P.L. Vornberger (1994). "Detailed elevation map of Ice Stream C, Antarctica, using satellite imagery and airborne radar," *Ann. Glaciol.*, 20, 327-335.

Bindshadler, P., E.P. Roberts, and A. Iken (1990). "Age of Crary Ice Rise, Antarctica, determined from temperature-depth profiles," *Ann. Glaciol.* 14, 13-16.

Bindshadler, P., S.N. Stephenson, E.P. Roberts, D.R. MacAyeal, and D.R. Lindstrom (1988). "Data report for the Siple Coast Project," NASA Technical Memorandum 10070.

Bliss, N.B. and L.M. Olsen (1996). "Development of a 30-arc-second digital elevation model of South America," paper presented at Pecora Thirteen Human Interactions with the Environment-Perspective from Space, Sioux Falls, South Dakota, August 20-22.

Bolstad, P. and T. Stowe (1994). "An Evaluation of DEM Accuracy: Elevation, Slope, and Aspect." *Photogrammetric Engineering & Remote Sensing* 60(11), 1327-1332.

Brassel, K. E. and R. Weibel (1988). "A review of conceptual framework of automated map generalization," *Int. J. of Geographic Information Systems* 2(3): 229-244.

Brenner, A. C., R.A. Bindschadler, R.H. Thomas, and H. J. Zwally (1983). "Slope-induced errors in radar altimetry over continental ice sheets," J. Geophys. Res., v. 88, p. 1617-1623.

Brooks, M.J. and B.K.P. Horn (1985). "Shape and source from shading," Proceedings of the International Joint Conference on Artificial Intelligence, Los Angeles, CA, pp. 932-936.

Burrough, P. A. (1986). *Principles of geographical information system for land resources assessment*, Clarendon Press, Oxford.

Burt, J.E., and G.M. Barber (1996). *Elementary Statistics for Geographers* (Second Edition, New York: Guilford Press).

Carrara, A., G. Bitelli, and R. Carla (1997). "Comparison of techniques for generating digital terrain models from contour lines," *Int. J. Geographical Information Science*, Vol. 11, No.5, 451-473.

Chen, L. C. and J. Y. Rau (1993). "A unified solution for digital terrain model and orthoimage generation from SPOT stereopairs." *IEEE Trans. Geosci. Remote Sensing* 31(6): 1243-1252.

Claus, M. (1984). "Digital terrain models through digital stereo correlation." *ISPRS Journal of Photogrammetry and Remote Sensing* 39(4-6): 183.

Cooper, A. P. R. (1994). "A simple shape-from-shading algorithm applied to images of ice-covered terrain." *IEEE Trans. Geosci. Remote Sensing* 32(6): 1196-1198.

Cooper, A. P. R., J.W. Thomson and E.M. Edwards (1993) "An Antarctic GIS: the first step," *GIS Europe*, 2(6), 26-28.

Cressie, N. (1993). *Statistics for Spatial Data*, Revised Edition, John Wiley & Sons, Inc.

Csatho, B., R.H. Thomas, W.B. Krabill (1996). "Mapping ice sheet topography with laser altimetry in Greenland," BPRC Technical Report No. 90-01, Byrd Polar Research Center, The Ohio State University, Columbus, Ohio, 53p.

Curlander, J. C. and R. N. McDonough (1991). *Synthetic Aperture Radar Systems and Signal Processing*. John Wiley & Sons, Inc.

Day, T. and J. P. A. L. Muller (1988). "Quality assessment of digital elevation models produced by automatic stereomatchers from SPOT image pairs," *Photogrammetric Record* XII(72): 797-808.

Declercq, F.A.N. (1996). "Interpolation methods for scattered sample data: accuracy, spatial patterns, processing time," *Cartography and Geographic Information Systems*, Vol. 23, No. 3, p128-144.

Deutsch, C. V. and A. G. Journel (1992). *GSLIB-Geostatistical Software Library and User's Guide*. Oxford University Press, New York.

Dowdswell, J.A. and N.F. McIntyre (1987). "The surface topography of large ice masses from Landsat imagery," *J. Glaciol.*, 33(113), 16-23.

Dowman, I. (1991). "A digital stereo workstation for SAR data," 1991 ACSM-ASPRS Annual Convention, Baltimore.

Dowman, I. J., H. Ebner (1992). "Overview of European Developments in Digital photogrammetric Workstations." *Photogrammetric Engineering & Remote Sensing* 58(1): 51-56.

Drewry, D., (1983). *Antarctica: Glaciological and Geophysical Folio*, Scott Polar Research Institute, University of Cambridge, Cambridge, England.

Durrant-Whyte, H.F. (1988). "Sensor models and multisensor integration," *International Journal of Robotics Research*, Vol. 7, No. 6, p97-113.

Ehlers, M., G. Edwards, and Y. Bedard (1989). "Integration of remote sensing with geographic information systems: a necessary evolution," *Photogrammetric Engineering & Remote Sensing*, v. 55, p. 1619-1627.

Eklundh, L. and U. Martensson (1995). "Rapid generation of digital elevation models from topographic maps," *Int. J. Geographical Information Systems*, Vol. 9, No.3, p329-340.

Ekoholm, S. (1996). "A full coverage, high-resolution, topographic model of Greenland computed from a variety of digital elevation data," *J. Geophys. Res.*, 101, 21,961-21,972.

ESRI (1991a). *Surface Modeling with TIN: surface analysis and display*, ESRI, Inc.

ESRI (1991b). *Cell-based Modeling with Grid -ARC/INFO User's Guide*. ESRI, Inc.

Ferrigno, J.G., J.L. Mullins, J. Stapleton, P. S. Chavez, Jr., M. G. Velasco, R. S. Williams, Jr., G. F. Delinski, Jr., and D. Lear (1996). Satellite Image Map of Antarctica, prepared by the USGS with joint support from the National Science Foundation Interagency Agreement OPP-9114787.

Fiksel, T., R. Harmann (1993). "ERS-1 data for mapping of the Antarctic Peninsula by shape-from-shading technique." Proceedings Second ERS-1 Symposium - Space at the Service of Our Environment, Germany, 11-14 Oct. 153-158.

Forster, R. R. K. C. Jezek, H. Sohn, A. L. Gray, and K. E. Matter. (1998). "Analysis of Glacier Flow Dynamics from Preliminary Radarsat InSAR Data of the Antarctic Mapping Mission," BPRC Technical Report 98-02, Byrd Polar Research Center, The Ohio State University, Columbus, Ohio.

Franke, R. (1982). "Smooth interpolation of scattered data by local thin plate splines." *Comp. & maths. with Appls.* 8(4): 237-281.

Franke, R. (1985). "Thin plate spline with tension," *Comp. Aided Geom. Design.* V.2, p87-95.

Frankot, R. T. and R. Chellappa (1987). "Application of a shape from shading technique to synthetic aperture radar imagery," Proceedings of IGARSS'87.

Frankot, R. T. and R. Chellappa (1989). "A method for enforcing integrability in shape from shading algorithms," in B.K. Horn and M.J. Brooks (eds) *Shape from Shading*, The MIT Press, Cambridge, Massachusetts, p89-121.

Frederiksen, P. (1981). "Terrain analysis and accuracy prediction by means of the Fourier transformation," *ISPRS Journal of Photogram. and Remote Sens.*, 36(4), 145-157.

Gao, J. (1995). "Comparison of sampling schemes in constructing DTMs from topographic maps." *ITC Journal*, v. 1.

Geischar, L. L., C. R. Bentley and L. R. Whiting (1992). "An analysis of gravity measurements on the Ross Ice Shelf, Antarctica," *Antarct. Res. Ser.*, 57, 105-155.

Ghiglia, D. C. and M.D. Pritt (1998). *Two-Dimentional Phase Unwrapping: Theory, Algorithms, and Software*, John Wiley & Sons, Inc.

Giles, P. T. and S. E. Franklin (1996). "Comparison of derivative topographic surfaces of a DEM generated from stereoscopic SPOT images with field measurements," *Photogrammetric Engineering & Remote Sensing* 62(10): 1165-1171.

GLAS Science Team (1997). *Geoscience Laser Altimeter System GLAS Science Requirements*, Version 2.0, August, 1997.

Goldstein, R. M., T.P. Barnett, and H. A. Zebker (1989). "Remote sensing of ocean currents," *Science*, v. 246, p. 1282-1285.

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

Goodchild, M., R. Haining, and S. Wise (1992). "Integrating GIS and spatial data analysis: problems and possibilities," *Int. J. Geographical Information System*, v.5, p.407-423.

Guindon, B. (1990). "Development of a shape-from-shading technique for the extraction of topographic models from individual spaceborne SAR images," *IEEE Trans. Geosci. Remote Sensing*, v. 28, p. 654-661.

Guritz, R. (1998). "Automated digital elevation model (DEM) production using ERS SAR Tandem pairs," Proc. IEEE 1998 IGARSS'98 Int'l Geoscience and Remote Sensing Symposium, Seattle, WA, USA, July 6-10.

Haining, R. (1994). "Designing spatial data analysis modules for geographical information systems," in S. Fotheringham and P. Rogerson (eds), *Spatial analysis and GIS*, Taylor & Francis.

Hampel, F.R. (1974). "The influence curve and its role in robust estimation," *J. Ameri. Statist. Assn.*, 69, 383-393.

Heipke, C. (1992). "A global approach for least-squares image matching and surface reconstruction in object space," *Photogrammetric Engineering & Remote Sensing* 58(3): 317-323.

Heuvelink, G. B. M., (1998). *Error Propagation in Environmental Modeling with GIS.* London: Taylor & Francis.

Hoar, G. J. (1982). *Satellite Surveying: Theory, Geodesy, Map Projection*, Magnavox, Torrance, California, USA.

Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins (1994). *Global Positioning System: Theory and Practice*, Spinger-Verlag Wien, 354 p.

Holtzman, J. C., V.S. Frost, J.L. Abbott, and V.H. Kaupp (1978). "Radar image simulation," *IEEE Transaction on Geoscience Electronics*, v. GE-16, p. 296-303.

Horn, B. K. P. and M. J. Brooks (1986). "The variational approach to shape from shading," *Computer Gragraphic Image processing* 33: 174-208.

Horn, B.K.P. (1990). "Height and gradient from shading," *Int. J. Computer Vision* 5:1, p37-75.

Hunter, G. J. and M. F. Goodchild (1997). "Modeling the Uncertainty of Slope and Aspect Estimates Derived from Spatial Databases." *Geog. Analysis* 29(1), 35-49.

Hutchinson, M. F. (1988). "Calculation of hydrologically sound digital elevation models," Third International Symposium on Spatial Data Handling, Sydney, Australia.

Hutchinson, M.F. (1989). "A New Procedure for Gridding Elevation and Stream Line Data with Automatic Removal of Spurious Pits." *Journal of Hydrology* 106, 211-232.

Ihde, J., U. Schirmer, and F. Stefani (1995). "Ice Mass Balance Investigation of Antarctica by Satellite Radar Altimeter Data," paper presented at The European Conference on Satellite Altimetry, Las Palmas, November 20-22.

Isaaks, E., and R. Srivastava (1989). *An Introduction to Applied Geostatistics*. Oxford University Press, New York, NY.

Jensen, J. R. (1980). "Stereoscopic statistical maps," *The American Cartographer*, Vol. 7, No. 1, pp25-37.

Jensen, J. R. (1995). "Issues involving the creation of digital elevation models and terrain corrected orthoimagery using soft-copy photogrammetry." *Geocarto International* 10(1): 5-21.

Jenson, S.K. and J.O. Domingue (1988). "Extracting topographic structure from digital elevation data for geographic information system analysis," *Photogrammetric Engineering & Remote Sensing* **54**, 1593-1600.

Jezek, K.C. (1999). RADARSAT Antarctic Mapping Project 1998-99 Progress Report," Byrd Polar Research Center, The Ohio State University, Columbus, Ohio.

Jezek, K.C. (1998). RADARSAT Antarctic Mapping Project. *BPRC Technical Report No. 17*, ISSN: 0896-2472, BPRC, The Ohio State University, Columbus, Ohio, 1998.

Joughin, I., D. Winebrenner, M. Fahnestock, R. Kwok, W. Frabill (1996). "Measurement of ice-sheet topography using satellite-radar interferometry," *J. Glaciol.*, Vol. 42, No. 140, 10-22.

Kaupp, V. H., L.C. Bridges, M.A. Pisaruck, H.C. Macdonald, and W.P. Waite. (1983). "Simulation of spaceborne stereo radar imagery: experimental results," *IEEE Trans. Geosci. Remote Sensing*, v. GE-21, p. 400-405.

Kiernan, R. (1998) "Ice sheet surface velocities along the Lambert Glacier Basin Traverse Route," *Antarctic CRC Research Report No.10* (in press), Australian Antarctic Division.

Kooij, M. W. A., B. Armour (1996). "A Workstation for spaceborne interferometric SAR Data," IGARSS'96 Remote Sensing for a Sustainable Future, Burnham Yates Conference Center, Lincoln, Nebraska, USA.

Kwok, R. and M. A. Fahnestock (1996). "Ice sheet motion and topography from radar interferometry," *IEEE Trans. Geosci. Remote Sensing*, Vol. 34, No.1, 189-200.

Lam, N.S. (1983). "Spatial data interpolation methods; A review," *American Cartographer*, 10(2), p129-149.

Leberl, F. G. (1990). *Radargrammetric Image Processing*, Artech House, Inc.

Leberl, F. G. (1998). "Radargrammetry," in F.M. Henderson and A. J. Lewis (ed) *Principles and Applications of Imaging Radar*, Manual of Remote Sensing, Third Edition, Vol.2, pp183-269, John Wiley & Sons, Inc.

Lee, J., P. K. Snyder (1992). "Modeling the Effect of Data Errors on Feature Extraction from Digital Elevation Models," *Photogrammetric Engineering & Remote Sensing* 58(10), 1461-1467.

Lemmens, M. J. P. M. (1988). "A survey on stereo matching techniques," International Archives for Photogrammetry and Remote Sensing, Kyoto.

Lin, Q., J.F. Vesecky, and H.A. Zebker (1994). "Comparison of elevation derived from INSAR data with DEM over large relief terrain," *Int. J. Remote Sensing*, v.15, p. 1775-1790.

Lingle, C. S., L. Lee, H. J. Zwally, and T. C. Seiss (1994). "Recent elevation increase on Lambert Glacier, Antarctica, from orbit cross-over analysis of satellite-radar altimetry," *Ann. Glaciol.*, 20, 26-32.

Liu, H. and K.C. Jezek, (1999). "Investigating DEM error pattern by directional variograms and Fourier Analysis," *Geographical Analysis*, Vol. 31, No. 3.

Liu, H. K.C. Jezek and M. O'Kelly (1999). "Automatic error detection and correction in irregularly spaced data sets," manuscript (in preparation).

Liu, H., K. C. Jezek and B. Li (1999) "Development of Antarctic digital elevation model by integrating cartographic and remotely sensed data: a GIS-based approach," submitted to Journal of Geophysical Research (accepted).

Lodwick, G. D. and S. H. Paine (1985). "A digital elevation model of the Barnes ice-cap derived from Landsat MSS data," *Photogrammetric Engineering & Remote Sensing* 51(12): 1937-1944.

Madsen, S.N. and H.A. Zebker (1998). "Imaging Radar Interferometry," in F.M. Henderson and A. J. Lewis (ed) *Principles and Applications of Imaging Radar*, Manual of Remote Sensing, Third Edition, V.2, pp359-380, John Wiley & Sons, Inc.

Mark, D. M. (1978). "Topographic properties of geographical surfaces: applications in cartography," Harvard Papers in Geographic Information Systems, Vol.5.

Marra, M., R. E. Carande, F. W. Leberl (1998). "Accuracy of DEMs from remotely sensed radar images," Proc. IEEE 1998 Int'l Geoscience and Remote Sensing Symposium, Seattle, WA, July.

Martin, T.V., J.H. Zwally, A.C. Brenner, and R.A. Bindschadler (1983). "Analysis and Retracking of Continental Ice Sheet Radar Altimeter Waveforms," *J. Geophys. Res.*, 88, 1608-1616.

Massonnet, D., and T. Rabaute (1993). "Radar interferometry: limits and potential," *IEEE Trans. Geosci. Remote Sensing*, 31, 455-464.

McDonald, R.A. (1995). "Opening the Cold War Sky to the Public: Declassifying Satellite Reconnaissance Imagery," *Photogrammetric Engineering & Remote Sensing*, 61 (4), 380-390.

McIntyre, N. (1991). "Mapping ice sheets with the altimeter." *Int. J. Remote Sensing* 12(8): 1775-1793.

Miller, S. B., U. V. Helava (1992). "Softcopy photogrammetric Workstations." *Photogrammetric Engineering & Remote Sensing* 58(1): 77-83.

Mitas, L., and H. Mitasova (1988). "General variational approach to the interpolation prblem," *Comp. Math. Appl.*, V.16, p983-992.

Mitasova, H. and J. Hofierka (1993). "Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry analysis," *Math. Geology.*, V. 25, p657-669.

Mitasova, H. and L. Mitas (1993). "Interpolation by regularized spline with tension: I. Theory and implementation," *Math. Geology*, Vol. 25, No. 6, p641-655.

Monckton, C. G. (1994). "An Investigation into the Spatial Structure of Error in Digital Elevation Data," In M. Worboys (ed), *Innovations in GIS* 1, pp. 201-211. London: Taylor and Francis.

Mrstik, V., G. VanBlaricum (1996). "Terrain height measurement accuracy of interferometric aperture radars." *IEEE Trans. Geosci. Remote Sensing* 34(1): 219-228.

Muhleman, O.D. (1964). "Radar scattering from Venus and the Moon," *Astronomical Journal*, Vol. 69, pp.34-41.

O'Kelly, M. (1994). "Spatial analysis and GIS," in Fotheringham, S., and Rogerson, P., editors, Spatial analysis and GIS, Taylor & Francis.

Oswald, H. and H. Raetzsch (1984). "A System for Generation and Display of Digital Elevation Models." *Geo-Processing* 2, 197-218.

Pardo-Iguzquiza, E., M. Chica-Olmo (1993). "The Fourier Integral Method: An Efficient Spectral Method for Simulation of Random Fields." *Math. Geol.* 25(2), 177-217.

Parkinson, B. W., J. J. Spilker, P. Axelrad, and P. Enge (1996). *Global Positioning Sytsem: Theory and Applications*, Washington, American Institute of Aeronautics and Astronautics, Inc.

Partington, K. C., W. Cudlip (1991). "An assessment of the capability of the satellite radar altimeter for measuring ice sheet topographic change." *International Journal of Remote Sensing* 12(3): 585-609.

Pentland, A.P. (1990). "Linear shape from shading," *Int. J. Computer Vision*, 4, 153-162.

Phillips, H.A., I. Allison, M. Craven, K. Krebs and P. Morgan (1996). "Ice velocity, mass flux and grounding line location on the Lambert Glacier-Amery Ice Shelf System, Antarctica," paper presented at AGU 1996 Western Pacific Geophysics Meeting, Brisbane, Australia, July 23-27.

Prati, C. and F. Rocca (1990). "Limits to the resolution of elevation maps from stereo SAR images." *Int. J. Remote Sensing* 11(12): 2215-2235.

Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (1992). *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edition, Cambridge University Press.

Rapp, R.H. (1991) "Geoid determination," in D.H. Elliot, W. Strange and I .M. Willians (ed) *Report of a workshop: utilization of the Global Positioning System (GPS) in addressing scientific problems in Antarctica, BPRC Technical Report Number 91-02*, P. 13, Byrd Polar Research Center, the Ohio State University.

Rapp, R.H., Y.M. Wang and N.K. Pavlis (1991). "The Ohio State 1991 Geopotential and Sea Surface Topography Harmonic Coefficient Models," *Rep. 410*, Dept. of Geodetic Science and Surveying, The Ohio State University.

Retzlaff, R., N. Lord and C.R. Bentley (1993). "Airborne-radar studies: Ice Streams A, B, and C, west Antarctica," *J. Glaciol.*, 39, 495-506.

Ridley, J. K., S. Laxon, C.G. Rapley and D. Mantripp (1993). "Antarctic ice sheet topography mapped with the ERS-1 radar altimeter." *Int. J. Remote Sensing* 14(9): 1649-1650.

Robinson, G.J. (1994). "The Accuracy of Digital Elevation Models Derived from Digitized Contour Data." *Photogrammetric Records* 14(83), 805-814.

Rodriguez, E. and J.M. Martin (1992). "Theory and design of interferometric synthetic aperture radars," *IEEE Proceedings-F*, Vol. 139, No.2., 147-159.

Scambos, T.A. and M.A. Fahnestock (1998). "Improving digital elevation models over ice sheets using AVHRR-based photoclinometry," *J. Glaciol.*, V.44, No.146, 97-103.

Schreier, G. (1993). *SAR Geocoding: Data and Systems*, Kasrlsruhe: Wichmann.

Schutz, B.E. (1998). "Spaceborne laser altimetry: 2001 and beyond," in H.P. Plag (ed), Book of Extended Abstracts, WEGENER-98, Norwegian mapping Authority, Honefoss, Norway.

Small, D., C. Werner, and D. Nuesch (1995). "Geocoding and validation of ERS-1 InSAR-derived digital elevation models," *EARSeL Advances in Remote Sensing*, Vol. 4, No.2, 26-39.

Smith, D.A. (1996). "On the improvement of gravimetric geoid undulations and oceanographic dynamic topography," Report No. 436, Dept of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio.

Sohn, H., K. Jezek, F. Baumgartner, R. Forster and E. Thompson (1999). "Radar backscatter measurements from Radar SAR imagery of South Pole Station, Antarctica," Submitted to IEEE IGARSS'99 Int. Geosci. Remote Sensing Symposium.

Swithinbank, C. (1988). Antarctic: Satellite Image Atlas of Glaciers of the World, United States Geological Survey, Professional Paper 1386-B.

Thapa, K. and Bossler, J. (1992). "Accuracy of spatial data used in geographic information systems," *Photogrammetric Engineering & Remote Sensing*, 58(6), 835-841.

Thomas, J., W. Kober, and F. Leberl (1991). "Multiple image SAR shape-from-shading," *Photogrammetric Engineering & Remote Sensing* 57(1): 51-59.

Tomlin, C. D. (1990). *Geographic Information System and Cartographic Modeling*, Prentice Hall.

Ulaby, F.T., R.K. Moore, A.K. Fung (1982). *Microwave Remote Sensing*, Vol. II, Artech House, Inc.

Usery, E.L. (1991). "Stereo display techniques for GIS," 1991 ACSM-ASPRS Annual Convention, Vol.4 GIS, pp.227-236, Baltimore, USA.

Vincent, R. K., M.A. True, and P.K. Pleitner (1987). "Automated extraction of high resolution elevation data from SPOT stereo images," Proceedings SPOT 1: Image Utilization, Assessments, Results: Paris, France, p. 1339-1346.

Warntz, W. (1966). "The topology of socio-economic terrain and spatial flows," Papers of the Regional Science Association 17: 47-61.

Warren, S.G. (1982). "Optical properties of snow," *Rev. Geophys. Space Phys.*, 20(1), 67-89.

Watson, D. F. (1992). *Contouring: A guide to the Analysis and Display of Spatial Data.* Oxford; New York. Pergamon Press.

Weaver, H. J. (1983). *Applications of Discrete and Continuous Fourier Analysis*, John Wiley and Sons, New York.

Weber, D. and E. Englund (1992). "Evaluation and comparison of spatial interpolators," *Math. Geol.*, Vol. 24, No.4, p381-391.

Weibel, R. (1992). "Models and experiments for adaptive computer-assisted terrain generalization," *Cartography and Geographic Information Systems* 19(3): 133-153.

Weibel, R. and M. Heller (1992). "Digital terrain modelling," in D. Maguire, M.F. Goodchild and D. W. Rhind (eds) Geographical Information Systems: principles and applications. Longman, London, pp. 269-297, Vol. 1.

Wilcox and H. Moellering (1995). "Pass location to facilitate the direct extraction of Warntz networks from grid digital elevation models," Proceedings of Auto-Carto 12.

Wildey, R.L. (1975). "Generalized photoclinometry for Mariner 9," *Icarus* 25:613-626.

Wood, J. and P. Fisher (1993). "Assessing interpolation accuracy in elevation models," *IEEE Computer Graphics & Applications*, 13(2): 48-56.

Yoeli, P. (1984). "Cartographic Contouring with Computer and Plotter." *The American Cartographer* 11, 139-155.

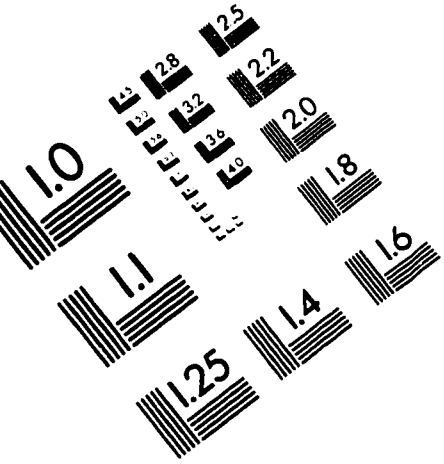
Zebker, H. A. and R. M. Goldstein (1986). "Topographic mapping from interferometric synthetic aperture radar observations." *J. Geophys. Res.* 91(b5): 4993-4999.

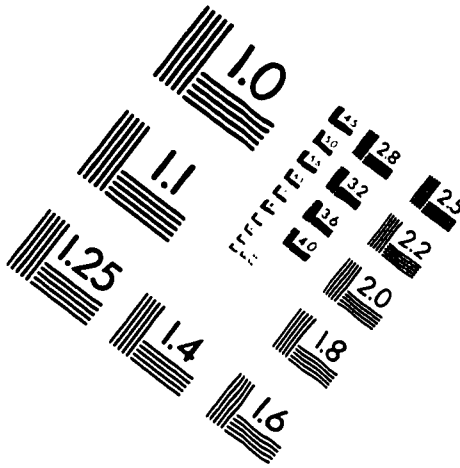
Zebker, H. A., C.L. Werner, P.A. Rosen, and S. Hensley (1994). "Accuracy of topographic maps derived from ERS-1 interferometric radar," *IEEE Trans. Geosci. Remote Sensing*, v. 32, p. 823-836.

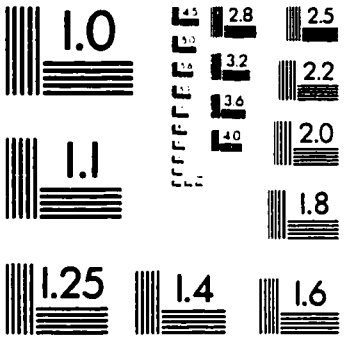
Zwally, H. J., A. C. Brenner, J. DiMarzio, and M. Giovinetto (1997). "Ice Sheet Topography, Slopes, and Flow Directions from ERS Altimetery," (abstract), *The 3rd ERS Symposium (ESA)*, 18-21 March 1997 Florence (Italy).

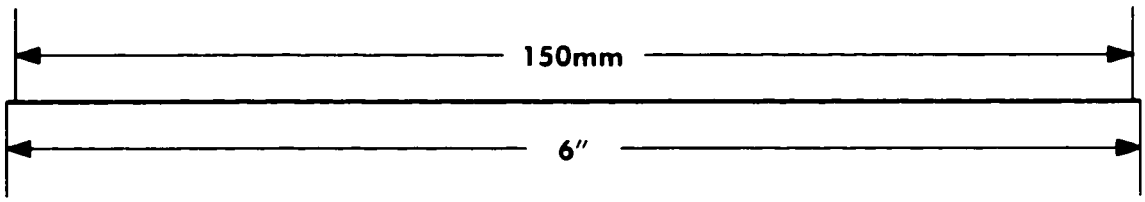
Zwally, H. J., R.A. Bindshadler, A. C. Brenner, T. V. Martin, and R.H. Thomas (1983). "Surface elevation contours of Greenland and Antarctic ice sheets," *J. Geophys. Res.*, 88, 1589-1598.

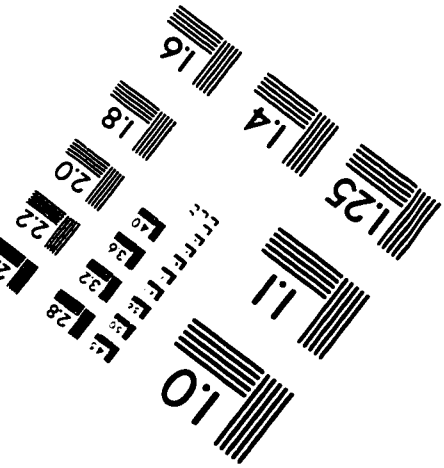
IMAGE EVALUATION
TEST TARGET (QA-3)

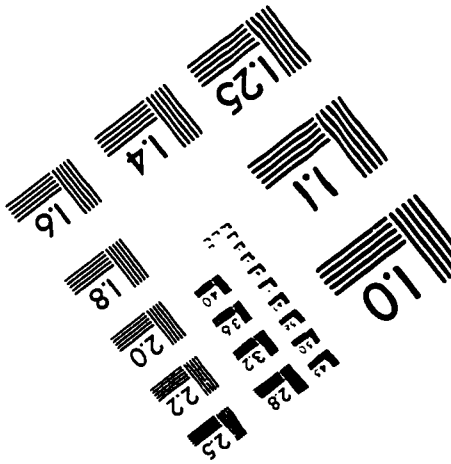












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved