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

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

Order Number 9416061

Automatic extraction of terrain features from digital terrain data: A multi-faceted study

Lay, Jinn-Guey, Ph.D.

University of Hawaii, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

AUTOMATIC EXTRACTION OF TERRAIN FEATURES FROM DIGITAL TERRAIN DATA: A MULTI-FACETED STUDY

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

GEOGRAPHY

DECEMBER 1993

By

Jinn-Guey Lay

Dissertation Committee:

Everett A. Wingert, Chairperson
Thomas W. Giambelluca
Matthew P. McGranaghan
Lyndon L. Wester
Russell S. Yost

ACKNOWLEDGEMENTS

I am grateful to many researchers and colleagues for their helps and supports. Thanks to Jay Lee, who has discussed the generation of TINs with me and sent me a copy of the source codes for TINs generation. The program for computing zero-crossings from Ci-Xiang Zhan is gratefully received. I owe Qi-Ming Huang a special appreciation for several sleepless nights of proof-reading. Many colleagues in the Geography Department also deserve credits for helping me in one way or the others; to Jonathan Chow, Nancy Elmer, Yuan-Qing Li, Richard Stone, and Eric Yamashita, I thank you.

To my wife Mei-Chai and two children, Yu-An and Yu-Han, I treasure the times we shared together in Honolulu with both happiness and difficulty . To my parents, I love you for the constant encouragement and support.

ABSTRACT

Ridge and valley lines are important terrain features to many scientific endeavors and practical applications. They are extracted manually from topographic maps traditionally. Besides being tedious, the manual process involves much arbitrariness by human interpreters and the results are not repeatable nor consistent. The lack of repeatability and consistence undermines the usefulness of the extracted results. The increasing availability of digital terrain data provides an alternative that may remedy the shortcomings of manual extraction.

The automatic delineation of terrain features is a multi-faceted task that is relevant to cognitive issues, terrain modeling, and computer implementation. This research has identified four groups of methods for the extraction of ridge and valley lines. They are: symbolic approach, tracing approach, profiling approach, and hydrological approach. The embedded meaning of ridge and valley lines in each of these methods is investigated. A series of tests are conducted inside computers to evaluate the performance of these methods. A primary investigation on the symbolic method concludes that the difficulty pertaining to the generation of TINs undermines the feasibility of the symbolic approach, thus it is not pursued further. The other three groups of methods have been tested and compared on the basis of the numbers, continuity, agreement, and positional accuracy of the extracted features. It is concluded that the hydrological approach performs the best generally.

The hydrological approach, instead of emulating manual interpretation in computers, takes an innovative approach that makes good use of the computational power of modern computers. By defining ridge and valley lines with accumulation values, this method is less sensitive to local terrain variations and extracts a rather continuous and complete result. In contrast, the tracing and profiling methods attempt to emulate manual process in extraction and the outcomes turn out to be not satisfactory. The various performances of the three methods present a notion that direct replication of human knowledge into computers is not necessarily feasible in the development of automatic methods. Several topics for future research are identified and subject to further study.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter 1. Introduction

1.1 Importance of Terrain Analysis

Terrain features such as plains, rivers, mountains, and rills exert a great influence upon human activity on the earth's surface. Natural catastrophes caused by volcanic eruptions, floods, and landslides are either processes of terrain evolution or the consequence of certain terrain conditions. Analysis of terrain may help to anticipate and so reduce the damage caused by these hazards. The terrain exerts a great influence on the stability of the land intended for the construction of bridges and mountain highways. Soil conservation is important to sustain agriculture and water supply. The erodibility of soil is greatly affected by the steepness and length of slope, which are factors of terrain. These examples portray the wide relevance and importance of terrain to human beings. For both practical applications and intellectual curiosity, terrain is studied by researchers from a number of fields including physical geography, geology, soil science, and civil engineering. By integrating terrain analysis with design and management projects, e.g., site selection, environmental impact assessment, and water resource management, researchers can provide better policies for maintaining the physical environment.

The approaches to terrain analysis vary from one discipline to another. There is no universal definition of terrain analysis and how it is being conducted. In general, terrain analysts investigate the morphometry, spatial difference, and the evolution of terrain. Morphometry deals with the shape, form, and pattern of the ground surface, such as slope, aspect, drainage pattern, drainage density, and bifurcation ratio: the ratio of the number of stream segments of one order to the number of stream segments of the next higher order. These factors are often measured or represented quantitatively. Experienced researchers interpret these values to gain understanding of an area, and apply them to further investigation of its terrain. The differences in the spatial aspects of terrain are of interest to some geomorphologists. The identification and explanation of such differences constitute a major component of geographic analysis by geomorphologists. One aspect of the study of terrain processes aims to identify the cause-and-effect relation of terrain evolution, such as erosion and mass movement. The dynamic analysis of terrain assists geomorphologists to fulfill one of the ultimate goals of scientific research, that is the prediction of change based on the current situation.

1.2 Feature Extraction and Terrain Analysis

Terrain features can be identified from their geometry (shape, size, and volume) or location and connection to other features. Terrain features can be large (i.e., glaciers, valleys, ridges, mountains); or small (i.e., rills and gullies). Ridge and valley lines define the character of terrain and play an important role in almost every type of terrain analysis.

From a morphometric perspective, ridges are important because they separate slope units and delineate water basins. The pattern of ridge lines can be used as a measure of the complexity of terrain. Valley lines, which dominate the pattern of drainage networks, are often thought of as the duality of ridges. Many morphometric measures are derived from the pattern of drainage networks; these include drainage density, stream ordering, and bifurcation ratio. Drainage patterns may reveal information concerning the parent rock and soil materials on the ground (Way, 1978, p.49). For example, a dendritic pattern of drainage network often indicates homogeneous, uniform soil and rock material, such as soft sedimentary rocks or volcanic tuff. The density and pattern of drainage and ridges networks can serve as indicators of areal difference. Furthermore, ridge and valley lines often serve as natural boundaries between different terrain units. The ridge crest of the Koolau Range on the island of Oahu, for example, separates the windward side from the leeward side. Climatologically, the ridge is a natural barrier intercepting the north-east trade wind, and inducing heavy precipitation. The windward side tends to receive more precipitation than the leeward side. The difference in precipitation strongly affects the morphometry, soil, and vegetation between these two sides of the range. These spatial patterns are of interest to many fields of study. Stream channels are the vessels of the earth. They carry water as well as sediments from mountains to flood plains and then to the ocean. For a dynamic analysis of terrain, streams and rivers are a major force in carving and shaping of the earth's surface. The density and pattern of drainage networks affect the hydrography of a water basin, such as a time-lag of storm flood, flow velocity, discharge, and frequency of floods. These are important factors affecting the evolution of terrain.

The focus and approach of terrain analysis may be different between individual research projects, yet the definition of terrain features from available data is almost always a preliminary step to any terrain analysis task and this is conventionally referred to as the "extraction of terrain features". This usage will be followed for the remains of the narrative. This extraction process used to be carried out by human interpreters using topographic maps. Automatic methods need to be developed to speed and systematize tasks that formerly had to be done manually.

1.3 Computers and Terrain Analysis

There are two specific reasons to use computers for terrain analysis. First, large quantities of terrain data are now being collected and stored in a digital form. The availability of digital data not only makes automatic methods simpler but the abundance of the data demands some sort of automatic method to handle it. Computers excel in dealing with fixed-format data when numerous computations are routinely required. It is difficult, if not impossible, for humans to deal with these data manually. The application of computers using automatic methods is desirable as well as necessary. Second, there is a growing appeal to adopt quantitative methods and numeric modeling in terrain analysis. Instead of descriptive and qualitative analysis of terrain, quantitative approaches are being used for the comparison and analysis of terrain (Evans, 1972; Hobson, 1972; Mark, 1975). There are increasing attempts to simulate natural phenomena with computer models and mathematical equations (Sprunt, 1972; Anderson *et al.*, 1988; Clarke, 1988). These research endeavors require a wide range of data in numeric format. Methods that provide information automatically as input to these models will facilitate this research. The hydrological model of a watershed is such an example. As slope is a major factor affecting hydrography it should be included in any such model. However, using traditional methods on topographic maps, the resulting slope data are normally given in categories (i.e., 10-20%), rather than as a discrete value for each individual unit area. Such categorization of slope values constitutes a loss or generalization of information and is undesirable for a quantitative model. With the use of digital data and automatic methods, slope information can be computed quickly and accurately in a format better suited to the modeling task.

There have been notable advancements in the application of computer techniques for terrain analysis during the past two decades. In the early 1970's, the concept of using a matrix of numbers to represent terrain was still an idea to be implemented (Evans, 1972). Today the use of digital elevation data is commonplace. With digital elevation data and computers, researchers within the geosciences and related fields have been able to analyze terrain more efficiently and precisely. New applications of computer processing are continuously being introduced. There is much room for further development, just as there are always questions to be solved.

1.4 Automatic Extraction of Terrain Features

Ridge and valley lines are traditionally delineated from topographic maps by interpreters based on the curvature and density of contour lines. The process of manual extraction is tedious. After extraction, to convert these valley and ridge data into a computer readable format, a digitizing process is needed, which is equally tedious. Given the fact that more terrain data are becoming available in digital format, the extraction process will be even more tedious and complicated for human if they have to work from large matrices of numbers. Since it is almost impossible for a human to work on the digital data directly, an extra conversion from digital data to a topographic map may be needed to continue using traditional method. An automatic method that can extract terrain features from digital terrain data will be highly desirable for this reason alone. Another major drawback of manual extraction is its arbitrariness, as shown by Mark (1983). There is no clear-cut rule of where a valley starts and where a ridge ends; thus, consistent delineation of valley and ridge networks is unlikely when the interpreters or the occasions vary. The following figures illustrate this problem. Figure 1.1a is a contour map extracted from a 1:24,000 topographic sheet of Kaneohe, Oahu. Figure 1.1b shows the stream channels of the same area from the same topographic sheet. The valley lines of this area have been extracted by two different interpreters and their results are illustrated in Figure 1.2. These two interpreters delineated the number and length of valley lines differently. In fact, it is common that the same interpreter may not be able to repeat his or her result.

(a) (b)

Figure 1.1: (a) A sample contour map; (b) its stream channels.

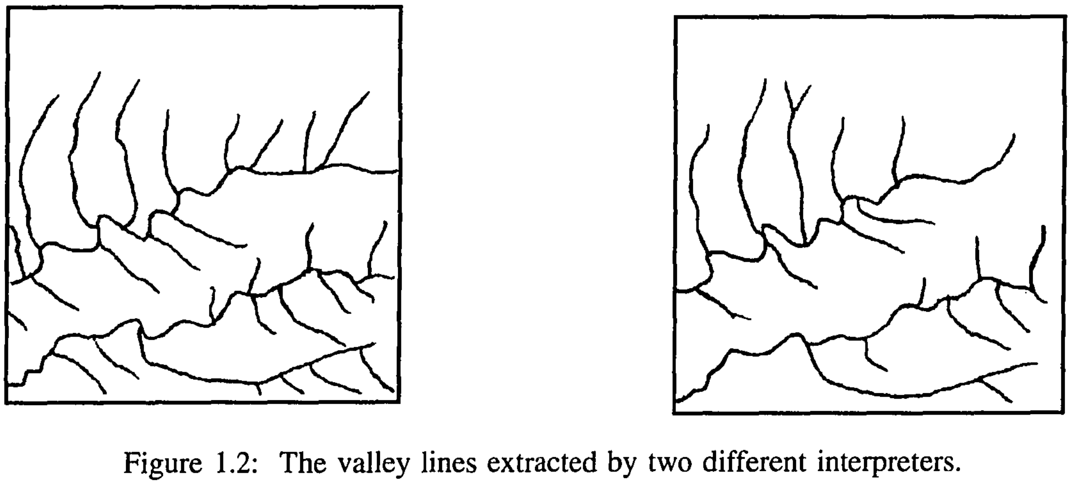


Figure 1.2: The valley lines extracted by two different interpreters.

Replication of the same result is a criterion of scientific experimentation, yet manual extraction is not repeatable. The lack of repeatability may impose problems when applications are developed from these variable extracted terrain features. A third order stream, for example, may otherwise be classified as a fourth order stream when the extent of a drainage network is extracted differently. Thus, the inconsistency of definition of stream order undermines the value of this method of analysis. In contrast to human interpretation, automatic extraction will always apply the same criterion to different areas or to the same area on repeated interpretations, and thus the terrain features can be consistently extracted. For practical reasons, an automatic method that extracts terrain features consistently will benefit researchers of many fields and is highly desirable.

Beyond practical applications, there are theoretical aspects pertaining to the development of automatic methods for feature extraction. An emerging branch of computer science, artificial intelligence (AI), emphasizes developing computer systems that are possible to perceive, reason, and act (Winston, 1992, p.5). AI technology has been applied in robotics, natural language processing, and expert systems with some successful examples. These intelligent computer systems, which work as human experts in specific domains, were once labeled as the next generation of computers. In geography, Robinson *et al.* (1986) identified four geographic tasks suitable for AI, one of which was automatic extraction of features from digital data. Many researchers have also been trying to develop intelligent computer system to extract geographic features, besides elevation data, from remotely sensed imagery (Erickson *et al.*, 1984; Goodenough *et al.*, 1987). A common challenge to the automatic extraction of features is to develop operational definitions of various features that are suitable to be used in computers. The development of operational definitions involves human cognition, the nature of digital data, and technical concerns in computer implementation. The problems facing automatic extraction of terrain features pertain to extraction of other objects as well. Thus, the research into methods of terrain feature extraction is a specialized case in the general development of automatic systems of feature extraction.

There has been much effort expended to develop automatic methods to extract terrain features from digital terrain models, yet many questions remain unanswered. Having many applications within various fields, and being a subject of great interest in its own right, the automatic extraction of valley and ridge lines has intrigued the author, who has made it the topic for this dissertation.

The extraction of terrain features is closely related to the terrain model being used. Since digital terrain modeling is relatively new, Chapter Two will introduce and investigate the nature of these models. Chapter Three will address the problems of feature extraction, review previous work, and then identify the research problem and goals of this study. Chapter Four explores the generic issues of *definition.* Chapters Five through Eight contain a sequence of testings and analyses. The last two chapters evaluate the performance of automatic extraction methods, identify a satisfactory method, and propose topics for future research.

Chapter 2. Digital Terrain Models

2.1 Introduction to Terrain Models

Models are widely used by scientists of various fields. As defined by Haggett, "a model is an idealized representation of the real world built in order to demonstrate certain of its properties" (Haggett, 1975, p.16). Most models are used for two main reasons. First, they are used to illustrate phenomena that are either too small or too huge for human eyes, for example, molecular structures of chemical compounds and planet systems. This type of model preserves the shape and pattern of the real-world phenomena at different scales and are called 'iconic models' by Haggett (1975, p.17) and Cromley (1992, p.6). Iconic models normally involve a limited degree of abstraction and can be directly seen and understood by most people. Second, models are used to highlight certain properties of complicated phenomena. As stated by Chorley and Haggett,

> Models can be viewed as selective approximations which, by elimination of incidental detail, allow some fundamental, relevant or interesting aspects of the real world to appear in some generalized form. Thus models can be thought of as selective pictures and 'a direct description of the logical characteristics of our knowledge of the external world shows that each of these pictures gives undue prominence to some features of our knowledge and obscures and distorts the other features that rival pictures emphasize....' (Chorley and Haggett, 1967, p.23).

This paragraph explains the generalization function built into most models. Complicated phenomena can be better understood or analyzed through selection, reduction, and distortion of original properties.

Many geographic models have been developed and adopted by researchers for various purposes. Because of the relevance of terrain to most geographic and environmental research, terrain models are the most widely used. A terrain model is any representation of the earth's surface that can help us comprehend or analyze terrain. The popular use of terrain models can be attributed to the two reasons mentioned in the previous section. Very often, researchers may find the terrain of their study area too large to comprehend when on the ground, especially where the terrain is rough. Thus, the adoption of a terrain model is needed in order to gain a whole view of a large area. Furthermore, much of the topographic detail on the earth's surface is often beyond the interest of research and needs to be generalized. The trivial features of terrain can be neglected on a model so that researchers can concentrate on more important properties of terrain.

A terrain model is an abstract and generalized form of the earth's surface. The degree of abstraction and generalization vary from one model to another. This variation affects the representativeness and usefulness of terrain models and, therefore, deserves our attention. The difference in the degree of abstraction and its implications can be illuminated by a comparison between a physical model and a contour map. A physical model represents the earth's surface explicitly with a three dimensional solid block, which can be easily comprehended by most people and thus is often presented in the information centers of national parks and so on. In contrast, a contour map represents the earth's surface more abstractly. The geometry of the ground is transferred into a pattern of contour lines, which is not as comprehensible to most people as physical models. However, the comprehensibility of these two models does not directly associate with their usefulness. Although physical models are easier to perceive, they are not very useful when certain measurements and analyses need to be made. For example, measuring the relief and steepness of terrain is difficult on a physical model. To the contrary, such morphometric properties can be easily measured from contour maps by experienced interpreters.

While the degree of abstraction may vary among different types of terrain models, the degree of generalization varies within terrain models of the same type. Contour maps of different scales, for example, show a different degree of generalization. The implication of different generalization to terrain modeling can be illustrated by the concept of fractal developed by Mandelbrot (1967). A fractal object is always inexact in its geometry and carries self-similarity, a property by which a subset, when magnified to the size of the whole, is indistinguishable from the whole (Goodchild and Mark, 1987; Clarke and Schweizer, 1991). Features with such property tend to reveal more detail with increasing resolution. Figure 2.1 shows a sample fractal feature in which the same shape appears in various levels. From another aspect, these lines can also be thought of as different generalization of a same fractal feature.

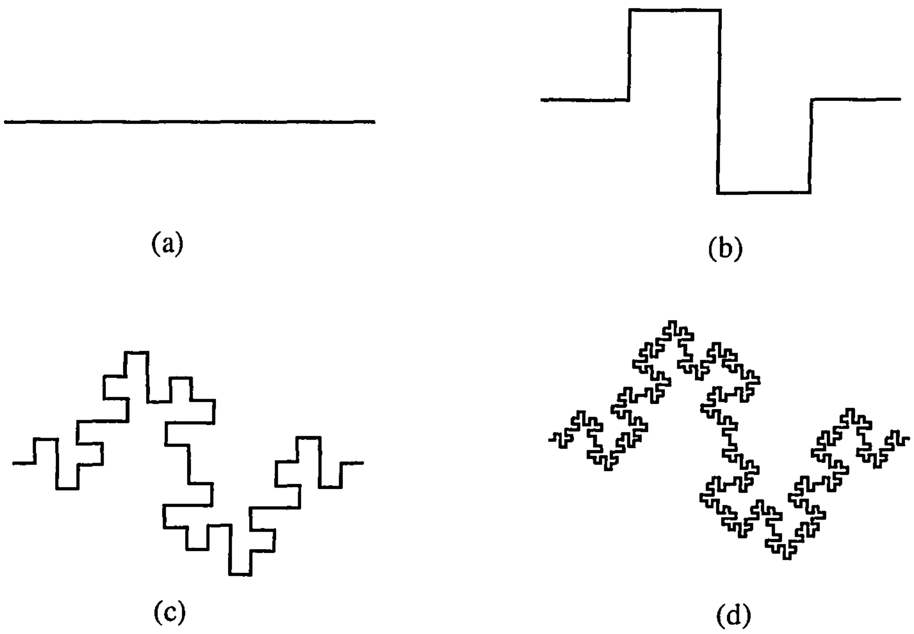


Figure 2.1: A sample of fractal plot (redrawn after Lam and De Cola, 1993).

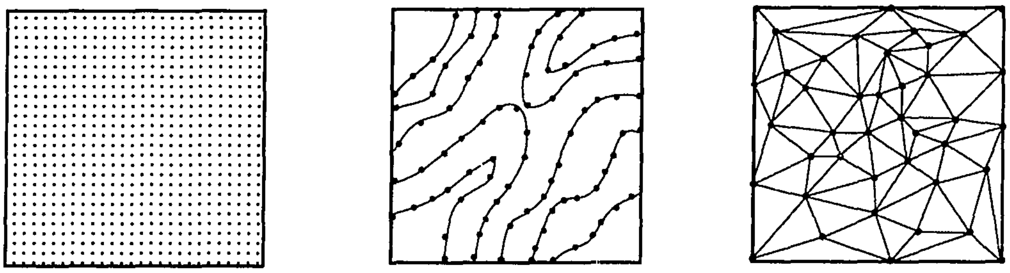
To some extent, terrain surface and its profiles are fractal by nature (Clarke and Schweizer, 1991) Therefore, the amount of detail of terrain depends on the resolution it is being studied. All terrain models are generalized forms of the earth's surface at arbitrary scales with selected details. In terrain analysis, researchers derive much information from terrain models, e.g., slope and aspect. The various generalization of terrain models affects measurement and analysis derived from these models. Any measurement analysis with no regard to scale and generalization is meaningless. Thus, an extensive understanding of the nature of terrain models is necessary for their proper use.

The research in this study deals with digital terrain models (DTMs), which are relatively new compared to traditional terrain models. The following section introduces the structure of DTMs and investigates their characteristics.

2.2 Digital Terrain Models

"A digital terrain model may be any numeric or digital representation of the elevations of all or part of a planetary surface, given as a function of geographic location" (Mark, 1984). There have been attempts to use mathematical equations to model terrain surfaces, but no practical application have been reported. Given the complex and variable nature of terrain, a feasible mathematical model for real-world application does not seem very likely. Currently, most digital terrain models explicitly record elevation of the earth's surface at a collection of points.

Three major data models are currently used in DTM storage. They are regular-grid, digital contour, and triangulated irregular network (TIN) (Figure 2.2). The apparent differences of data format and organization are superficial. A profound difference among these three structures, however, is how they generalize the earth's surface, eliminate noise, and selectively sample ground truth into digital models. These generalization and sampling processes determine the nature and amount of information presented in the DTM. More details of these three structures follow.



(a) (b) (c)

Figure 2.2: Three structures of DTMs, (a) regular-grid; (b) digital contour; (c) TIN.

2.2.1 Regular-Grid Structure

A regular-grid structure stores an array of elevation data, taken from a matrix of equally-spaced points, to represent the earth's surface. It is a projection of a regular grid on the terrain surface and a recorded elevation value at every intersection of the grid (Figure 2.3).

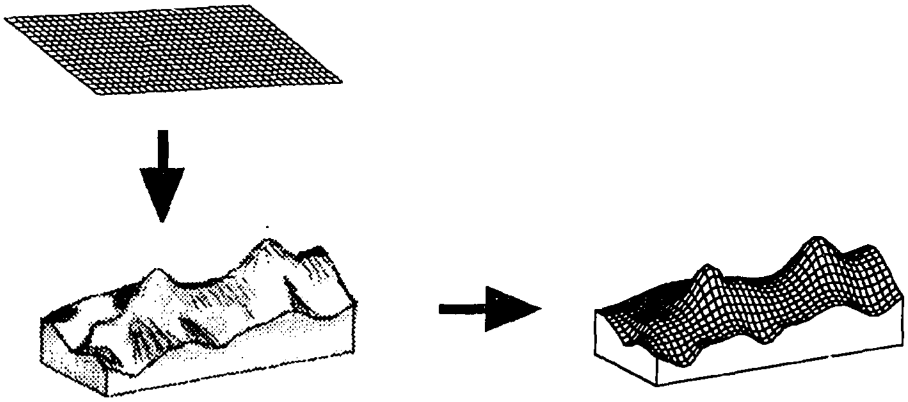


Figure 2.3: An illustration of a regular-grid DTM.

Elevation is measured at a fixed distance, normally referred to as data resolution, in orthogonal directions regardless of the complexity of terrain. Figure 2.4 is an example that generalizes the cross profile with a sequence of points, with variation between two adjacent points being ignored. Local relief with a horizontal extension less than the data resolution, such as (A) in Figure 2.4, will likely be ignored no matter how steep the relief is. On the other hand, a gentle wide relief as (B) will be recorded in the model. Horizontal extension of a relief is the major factor affecting the sampling process.

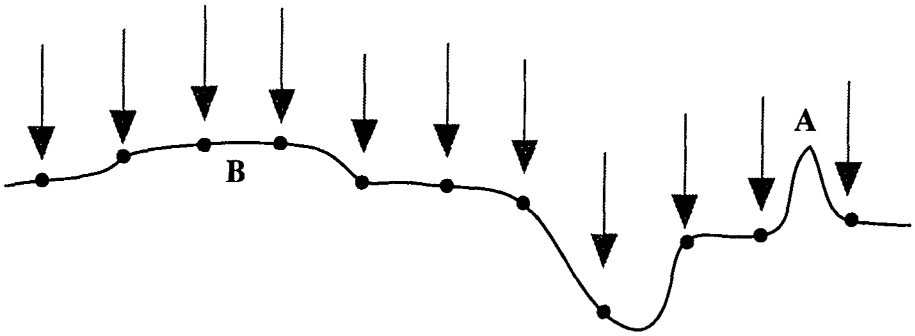


Figure 2.4: A cross-profile sampled by regular-grid.

The elevation of any location inside the data area can be referred to by a known elevation, within a maximum distance of 0.707 times the data resolution, which is one half of the diagonal distance between adjacent known points, as illustrated in Figure 2.5.

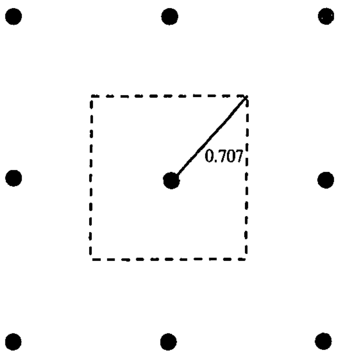


Figure 2.5: Distance within one pixel.

The average distance between any point and a known elevation can be derived from the mathematical equation:

$$d=\int\int\sqrt{x^2+y^2}/\Omega \qquad , \Omega : \text{area}$$

which is 0.383 times the data resolution. With a 30 meter resolution, these figures translate to 21.21 meters for maximum distance and 11.5 meters for average distance. There will be some variation of elevation within these distances, depending on the local complexity of terrain. A careful user of DTM should be aware of these variations.

The advantage of this structure is its simplicity in data format and consequently the ease in data retrieval and analysis. The U.S. Geological Survey (USGS) uses automatic photogrammetric systems to create digital elevation data in this format (USGS, 1983a). Several interpolation methods are also available to generate this regular- grid data from contour lines. This structure is currently the most available and frequently used structure of DTM.

2.2.2 Digital Contour

A contour is a smoothed line that passes through points having the same elevation. It is the line formed by the intersection of a level surface with the surface of ground. It was first introduced by marine cartographers to show the depth of water in the 1730's and the method was later applied to land (Harvey, 1980, p.182). Contour maps have been widely used to represent terrain relief since the second quarter of the nineteen century. On a contour map, the ground surface is generalized into a set of horizontal intersections sampled at a constant interval, as Figure 2.6 shows.

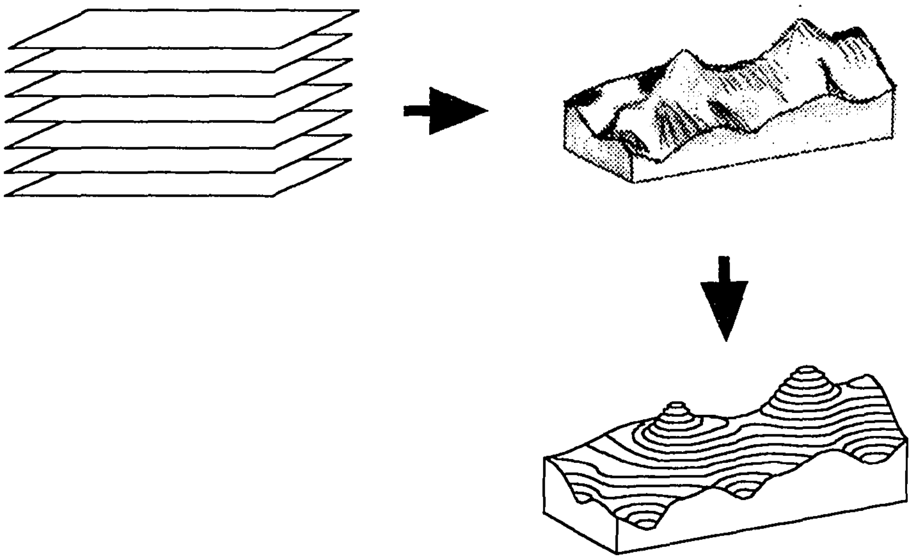


Figure 2.6: An illustration of contours.

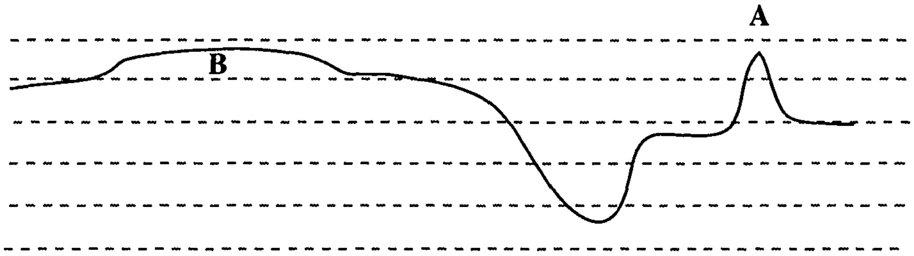


Figure 2.7: A cross-profile sampled by contours.

The cross profile in Figure 2.5 is sampled by contour as illustrated in Figure 2.7. The (A) feature will be recorded in the data because its relief is larger than the contour interval. On the other hand, feature (B) is being ignored because its relief is not large enough to cross the contour interval.

The elevation of every point within the map area can be estimated to a range, i.e., the elevations between two adjacent contour lines. The distance between an unknown point to its adjacent contour lines may be far, yet this estimate is valid regardless of the distance. Besides providing elevation data, contour maps also visually display the complexity of terrain by the varying density and shape of contour lines.

Contour maps are the major means for storing elevation data. By digitizing these contour maps, either through manual methods or scanning, digital contour data can be obtained. There are semi-automatic systems that compile contour lines from stereoscopic aerial photos. Although digital contour data can be obtained in many ways, they are seldom used directly in computer analysis. Contour lines are not easily manipulated by computers due to the inexplicit relation both between contours and data points of contours, thereby limiting their use in digital analysis.

2.2.3 Triangulated Irregular Network

A triangulated irregular network (TIN) is a system of triangulated facets to approximate the earth's surface (Peucker *et al.*, 1978). A TIN matrix consists of nodes, segments, and triangles. Nodes are points with elevation data; segments are non-intersecting linkages between adjacent nodes; triangles are the consequent facet of three adjacent segments. The construction of TIN consists of two steps: selection of points that will serve as nodes for networks and then connecting nodes into triangles. The selection of nodes is critical because nodes are the basic element of the TIN. The connection of nodes into networks is intricate. Different connections will render terrain differently. An arbitrary connection may fail to render the original terrain faithfully. The subtle nature of this structure introduces much uncertainty and difficulty in constructing a TIN. Building an accurate TIN is still a research topic attracting much interest (Lee, 1989, 1991; Tsai *et al.*, 1991).

One advantage of the TIN structure is its flexibility over different terrain areas. In a homogeneous terrain area where slope remains more or less unchanged, a small number of triangles will be sufficient to represent a large area closely. In an area where terrain is rugged and complex, the number of triangles may increase to reflect such complexity adequately.

2.3 Comparison and Conclusion

The volume of data is often an important concern for data usage. A larger volume requires more storage and more processing time. A regular-grid is rigid in data volume because the measurements are at a fixed distance regardless of the terrain complexity, with much redundancy in flat areas. Its data volume is roughly proportional to the size of data area. The data volumes of contour lines and TIN structures will vary with the complexity of terrain. A gentle terrain area requires fewer contour lines or TIN for rendering. The superficial redundancy and flexibility associated with each structure does not necessarily reflect on the efficiency of storage among different structures. With a regular-grid structure, the coordinates of each point are automatically indicated by their order and, therefore, need not be explicitly registered. Digital contours require explicit registration of x,y coordinates for each point on the lines. Triangulated irregular networks require definite x,y,z values for each node and also require extra storage to specify the connections between nodes.

Taking these into account, there is no apparent large advantage in storage size among these structures within the expected variation in terrain types.

Availability of data is another concern in adopting the structure of terrain models. There are many methods available to convert these three data structures from one to the other (Lam, 1983; Legates *et al.*, 1986; Lay, 1987). If necessary, conversion from one to the other is possible. However, conversion between different structures always causes loss of information and never makes the data more accurate. Unless there is no other choice, conversion should be avoided.

Currently, the regular-grid structure of DTMs is the most popular and most available format. In the United States, the United States Geological Survey produces and distributes regular-grid structure terrain models, named a Digital Elevation Model (DEM), at two different scales. In fact, in the United States, 'DEM' has a strong connotation with regular-grid DTM when used as a generic term. The series at 1:250,000 (USGS 1:250,000 DEM) scale is produced by interpolating digitized line data from 1:250,000 topographic maps. The coordinates are based on the longitude and latitude system and spaces between data points are 3 arc-seconds. The other series is parallel to 1:24,000 scale of topographic sheets, based on Universal Transverse Mercator (UTM) coordinates with a 30 meter resolution. Each file of this DEM series corresponds to a 1:24,000 or 1:25,000 topographic quadrangle map. These data are produced by either photogrammetric or interpolation methods. The Defense Mapping Agency (DMA) produces another series of data called digital terrain tapes (DTT), which are also distributed by the USGS. Each DTT file covers a 1 by 1 degree block, representing one half of a 1:250,000 scale map. The grid size of DTT is about 208 feet (USGS, 1983a).

The U.S. Geological Survey, one of the major producers of digital geographic data, does not publish digital data in TIN format. It may not be necessary to have such data since a TIN matrix is supposed to be flexible in presenting various details of terrain. TINs may be built from either regular-grid, contour data, or from field surveying anyway if needed. For large areas, field surveying is not very likely; conversion from DEMs or digital contours is the main approach of generating a TIN. Recently, the USGS has begun to release digital contour lines as part of the USGS digital line graphs (DLG) series. They, however, are not generally available. The regular-grid DEM is still the most widely available source of digital terrain models.

This chapter introduced the structure of DTMs and their characteristics. Many methods have been developed to extract terrain features automatically from digital terrain data. The nature of digital data will affect the extraction of terrain features. DTMs of regular-grid structure use *points* to represent surfaces; digital contours use *lines* to represent surfaces; TINs use *facets* to represent the surfaces. These different structures represent the original terrain differently. A sensible extraction method should take the unique characteristics of each structure into account.

Chapter 3. Research Questions and Purpose

3.1 Problem Domain of Terrain Feature Extraction

Systematically, the process of extracting terrain features can be divided into three steps: conceptualization, abstraction, and implementation. A brief review on the procedures of manual extraction is provided here for showing the rationale of these three steps. Because of their availability, contour maps are most often used in manual extraction. Before delineation of ridge and valley lines, interpreters using contour maps must form a concrete conceptual model of 'ridge' and 'valley' containing the characteristics of these objects. Such a task constitutes the *conceptualization* stage of work. Based on the conceptual model, interpreters transform these characteristics into the shape and pattern of contour lines, which involves a work of *abstraction.* Through this conversion, for example, a ridge is abstracted as a linear feature along all those points where contour lines show a convex down-hill curvature. Following the same manner, a valley line can be abstracted as a line along points where contour lines pose a down-hill concave shape. The pattern of contour lines becomes an operational definition of ridge and valley lines on a contour map. Based on such a definition, interpreters can delineate these terrain features by tracing across contour lines. This constitutes *implementation* work. The process of manual extraction has been well developed and widely adopted, thus anyone who can read contour maps can be trained easily to delineate valley and ridge lines following a set of rules without knowing the underlying conceptualization and abstraction processes. The development of automatic extraction is to construct such a set of rules to be implemented and executed in computers.

Computers differ from humans in their computation and reasoning ability. Contour maps and digital terrain data differ in their formats and characteristics. Due to such differences, those extraction procedures contained in manual extraction are not necessarily applicable to automatic extraction. The development of automatic extraction first requires a comprehensive review on the tasks involved. The methodology of automatic extraction may be different from that of manual extraction, while the necessity of the three steps are the same. This review starts with an explanation of these steps: conceptualization, abstraction, and implementation.

3.1.1 Conceptualization

People create and use various terms to organize their thinking and experiences. Gyorgy Kepes (Abler *et al.*, 1971, p.3) once wrote, "We make a map of our experience patterns, an inner model of the outer world, and we use this to organize our lives." Words and terms of human language are labels of these maps or models that represent our experience. Thus, each term carries a conceptual model that portrays its characteristics. We use these conceptual models in daily life whenever applicable, e.g., we call an object a 'mountain' when its characteristics fit our conceptual model of 'mountain'. Conceptual models exist and are used so natural that we are seldom aware of their existence.

The conceptual models in our mind are usually somewhat vague and difficult to describe. Most geographers, for example, will admit having used 'maps' for various purposes, yet what is a map? Topographic maps and road maps are some typical objects that all geographers will agree are maps. Yet, is an aerial photograph a map? How about a satellite image? Recent advances in the field of computer graphics and visualization are being introduced to display spatial information. Along with the diversity of the outlooks and materials of 'map', the definition of 'map' is getting both more complicated and more vague. The definition of 'map' is important and problematic enough to attract the research interest of Vasiliev *et al.* (1990). They investigated the nature of 'map' and identified five factors that contributed to the categorization of a map: location, graphic image, generalization, function, and prototype effect. However, the measurement of these factors to the categorization of map is often vague.

A clear description of our conceptual model for valley and ridge is essential before we can program computers to extract them from a digital terrain model. Therefore, an exploration for the characteristics of terrain features is needed for their extraction.

3.1.2 Abstraction

The delineation of terrain features from digital data is done by reference to models and not from the ground surface; therefore, it is necessary to associate our conceptual models with the representation of ground surface in various sorts of terrain models. This association should result in a solid and feasible definition, an operational definition of terrain features that can be applied to the specific form of a terrain model, i.e., terrain is represented differently by various terrain models, the operational definitions need to reflect model differences. On an aerial photo, ridges are associated with contrast between different tones while on a contour map they are associated with the convex pattern of contours. This example illustrates why we need to develop an individual operational definition to fit different data models.

The representation of terrain on digital terrain models can be very different from that on traditional models. The creation of new operational definitions based on the representativeness of these DTMs is needed in order to make the development of automatic extraction possible. The criteria for this operational definition are twofold: first, the definition must faithfully reflect the representation of terrain features on DTMs; second, the definition must be in a form that will enable computer processing. To extract valley lines from a regular-grid DTM, for instance, we need to first imagine how a valley line, based on our conceptual model, will be represented on this DTM, and then create a feasible computer definition. It is the author's view that this stage is the most critical in the entire process of feature extraction. How we define terrain features will greatly affect the results obtained.

3.1.3 Implementation

Although there are attempts to make smart computers that can learn by themselves, most current computers will solve problems only when the problem solving procedure is clearly defined. Linguistic definitions such as "convex" and "concave pattern" are not appropriate to computers since they are ambiguous and require subjective and arbitrary judgments, which are difficult for computers to execute. An operational definition to be used in computers needs to be precise and definable. Otherwise, the definition will be in vain if it is not in a form applicable to computers.

Given a feasible operational definition, the technical challenge is how to adapt it so that it works accurately, efficiently, and flexibly on a computer. Although the price of computer hardware keeps decreasing while performance keeps improving, efficient use of computer memory and the time needed for computations is still a concern to software development. Pragmatically, when the execution time of an algorithm is too great, it is simply not appropriate. Algorithm analysis aims to improve the efficiency of computation and storage and can make an otherwise too costly algorithm feasible. A flexible method that can work with different types of terrain and data sizes is preferable.

3.2 Introduction to Previous Research

The automatic extraction of terrain features has attracted much research attention in cartography and GIS, remote sensing, computer graphics, geomorphology, water resources, and geology in the past two decades. In a broad sense, the concept of terrain features has been applied in other fields of study as well. Warntz (1966) borrowed the concept of terrain features (pit, peak, pass, pale, and course line) and defined them in his study of socio-economic surface. Much research of the past two decades is presented in a methodological evolution from primitive to sophisticated techniques. There are also contemporary studies which developed in parallel with fundamentally different approaches and emphases. Such methodological variations often occur, not coincidentally, with variations in researchers' backgrounds. Researchers of different backgrounds often perceived terrain features differently. These different perceptions are reflected in the working definition of terrain features.

Over the years, noted contributions to this topic include studies by Peucker and Douglas (1975), Marks, Dozier, and Frew (1984), O'Callaghan and Mark (1984), Palmer (1984), Jenson (1985), Band (1986), Douglas (1986), Jenson and Dominigue (1988), Riazanoff, Cervelle, and Chorowicz (1988), Lammers and Band (1990), Smith, Zhan, and Gao (1990), and Chou (1992). Not all these studies, however, explicitly dealt with extraction of valleys and ridges. The research by Marks *et al.* delineates water basins from gridded DTM. Their method first sorts points in a descending order based on their elevation. The highest point is the head and sole boundary point of the first water basin being extracted. This method will then check the location of each next point in the descending list to see whether it is adjacent to the boundaries of any existing water basin. The point will be assigned to a particular water basin if it is adjacent to the basin, otherwise it will be assigned as the head of a new water basin. One major drawback of this method is that many grid points may have the same elevation which makes the comparison of elevation difficult. Also, this method only extracts water basins, not valley and ridge lines. The method developed by Chou (1992) delineates slope lines from digital contours stored in the ARC/INFO format. The vagueness of spatial relations between contour lines and between points of a contour introduces much tedious bookkeeping and computation to this method. As a result, this method is much more complicated and computation-intensive compared to the method of same function developed for a gridded DTM structure. Furthermore, Chou's implementation requires the user to specify the starting point of each slope line to be extracted and is not fully automatic.

The rest of the aforementioned studies present a rather complete procedure for the extraction of valley and ridge lines. Based on their methodology, this author categorized the rest of the studies into four groups: symbolic, profiling, tracing, and hydrological. Each group represents a unique perspective toward terrain feature extraction. The first approach is applied to triangulated irregular networks although the original researchers claimed their work was applicable to gridded DTMs as well. The remaining three approaches deal with gridded DTMs. Each of these groups are introduced below and subject to further tests and analyses in later chapters.

3.2.1 Symbolic Approach

Palmer (1984) used the Prolog language to explicitly define terrain features. Prolog, stands for *programming in logic* (Firebaugh, 1988, p.46), is a relatively new computer language that has a strong connection to the development of artificial intelligence technology. Unlike most conventional programming languages that specify problem-solving procedures step by step, Prolog language specifies rules or knowledge for problem-solving. Thus, it is very suitable to handle symbolic operations, a unique characteristic compared to many other computer languages that excel in numerical operations. With the expressive nature of Prolog in coding rules, the working definition of terrain features can be explicitly expressed in the Prolog language.

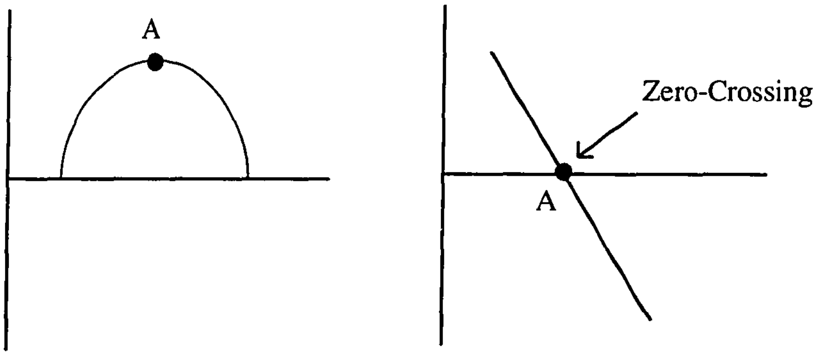
Much of Palmer's work emphasized the development of a formal definition of geographic features, rather than extraction of terrain features. Formal definitions were based on formal theories (e.g., Euclidian geometry) and consist of a small number of initial terms on which all other terms build. Formal definitions are considered alternatives used to avoid the problem of dependence on generally accepted terms or circular definitions which cause subjectivity and ambiguity. Following this study, Frank *et al.* (1986) examined the feasibility of creating formal definitions for physical geographic features. They concluded that research to analyze the problem of scale-dependency should be carried out first. The scale-dependency issue here refers to the amount of detail in a TIN. Such a conclusion is by no means an accident but rather an expectable consequence of the data structure they adopted. Their research adopted a TIN structure of DTMs for analysis, the only usage of TIN of all four reviewed groups. A TIN structure is flexible in data size and amount of detail. The amount of detail of the terrain data critically affects the number of features that can be extracted.

Therefore, the construction of TIN is a critical issue for the automatic extraction method from a TIN structure.

3.2.2 Profiling Approach

The shape of terrain profiles appears to be an obvious property of terrain features. Therefore, several algorithms have been developed by geographers and computer scientists separately based on this property (Peucker and Douglas, 1975; Haralick, 1983; Jenson, 1985). In this group of approaches, ridge lines are composed of points with convex slope inversion, and vice versa for valley lines. The HILO (high/low) algorithm proposed by Peucker and Douglas (1975) represents an early development of this approach. This method employs a local operator on a two row by two column area of DTM grid that flags the lowest point out of the four grid points. The local operator sequentially works on each 2 x 2 grid, from left to the right and top to bottom of the DTM area, repeating the flagging process. When the process completed, those points remaining un-flagged are considered ridge points. By flagging the highest point of each 2 x 2 window, this method can also identify drainage networks. Although this method works on 2 x 2 window superficially, the deep implication is in fact comparing the profiles of each 3 x 3 window of a DTM grid. A later development by Jenson (1985) explicitly compared the cross profiles of DTMs in each 3 columns by 3 rows of grid.

Zero-crossing, widely used in computer graphics for edge-detection, provides another approach to the profiling method. A zero-crossing can be defined as those points where the curve of a mathematical function passes the X-axis. In computer graphics, images are normally scanned and represented with a matrix of grey-level values for further operations. The data structure of these grey values is identical to that of gridded DTMs and digital remotely sensed imagery. Therefore, techniques developed for computer graphics and image processing are often applicable to analysis of gridded DTM data. A zero-crossing can be identified easily on a discrete data set such as gridded DTM. On a 3 x 3 window of a DTM grid, a center point that is either higher or lower than its two neighbors in any of the four cross profiles will be a zero-crossing (Figure 3.1). These zero-crossings are considered as part of ridge or valley line networks.



(a) (b)

Figure 3.1: An illustration of a zero-crossing, (a) the original curve; (b) its derivative.

Concerned with the noise and error in gridded DTMs, Haralick (1983) introduced a sophisticated method to find zero-crossings on such data. This method generates a two-variable cubic polynomial to fit elevations of each neighborhood of gridded DTMs. Ridge and valley lines are defined as those points close enough to zero-crossings of the first directional derivative taken in a direction that maximizes the second-directional derivative. Smith, Zhan, and Gao (1990) presented a two-step procedure for extracting valley lines built upon Haralick's method. The first step employs Haralick's method to extract pixels pertaining to valley lines from digital images. The second step is a tracing process that starts from particular extracted pixels of valley lines and delineate valley line networks of single-pixel width. The following review of the tracing approach will cover more details of this process.

3.2.3 Tracing Approach

To geologists, ridge and valley lines are continuous linear features that can be traced in several ways. Based on this concept, Riazanoff *et al.* (1988) developed three algorithms and named them as: streaming, walker, and main saddle points. These algorithms can extract both ridge and valley lines using different definitions. Each algorithm uses a 'selected point' and a 'progress constraint'. A selected point is one of those points described as singular, e.g., a saddle point, or a local maximum. These points are normally first identified during one pass of the whole DTM area. A progress constraint is a function that associates a candidate, its eight neighbors, parent (direction from which it comes from) and a series of children (directions of progress). A candidate is part of a feature if, and only if, it has children. Each child becomes a candidate in turn, and so on. Different selected points and progress constraints distinguish one tracing algorithm from the others. The tracing algorithms treat ridges and valleys as 'duals' of each other. By reversing the tracing criteria from highest to lowest, for example, an algorithm used to trace ridges is used to trace valleys.

Smith *et al.* (1990) integrated a tracing process as part of their extraction method. Their tracing approach started from the highest point of each valley line extracted by a zero-crossing approach. From each starting point, the progress constraint is to trace along the lowest neighbor until the tracing process encounters a pit, a point lower than all its neighbors.

3.2.4 Hydrological Approach

A hydrological approach was adopted by a group of researchers (O'Callaghan and Mark, 1984; Band, 1986; Jenson and Dominigue, 1988) that was concerned specifically with water flow. Early research focused on the extraction of drainage networks. Later developments proceeded to delineate watersheds and calculate stream ordering. O'Callaghan and Mark (1984) first presented this approach to extract a drainage network from a regular gridded DTM. Their method first smooths DTM data by a local averaging procedure that is designed to decrease noise characteristics of the DTM. Next, the program finds pits. The problem raising from pits, which will terminate a water flow, is solved by finding the overflow of a pit, or depressed area. An overflow is the point with lowest elevation along the boundary of a depressed area. Water flows that encounter a pit will be directed to the overflow point accordingly without being terminated. This approach then computes flow direction, the direction to one of its eight neighbors with lowest elevation, adjusted with distance, of each cell inside the data area. Next, the accumulative amount of input water is calculated for each cell based on flow direction. All those cells with input water in excess of a threshold value are marked as stream channels. By changing the threshold, users are able to extract stream channels of different hierarchies.

Some methodological improvements have been made in the hydrological approaches. Band (1986) incorporated a coding scheme to label each stream segment and water basin extracted. Lammers and Band (1990) later organized these extracted terrain features into an object-oriented geomorphometric database, which stores each segment and basin as an individual object rather than storing low-level raster data. Ehlschlaeger (1989) also contributed to the development of the hydrological approach by introducing a more efficient searching algorithm for tracing water flow. Automatic extraction of stream channels and water basins using this approach is now part of the GRASS (Geographical Resources Analysis Support System) package, a GIS software developed and made available to the public by the U.S. Army Construction Engineering Research Laboratory (U.S. Army CERL, 1992).

3.3 Research Questions and Purpose

Most previous studies focused on implementation and technical issues. Comparing these studies with the problem domain addressed in section 3.1, the author identifies a lack of investigation into conceptualization and abstraction issues. A complete investigation should incorporate all relevant issues of extraction in order to produce the most accurate result. Each method carries a definition of terrain features, explicitly specified as in Prolog or implicitly hidden in the algorithms. These definitions affect what and how terrain features will be extracted. This is an area critical to the development of automatic methods for terrain features extraction and deserves more research. This current research examines the performance of the four approaches through tests and analysis with a focus on issues relevant to operational definitions. The major research questions are listed below.

(1). What are human conceptual models for ridge and valley lines?

Conceptual models reflect our definition and specification of terrain features for various applications. To understand the definition of terrain features better, this research investigates issues relevant to human cognition and addresses the nature of conceptual models.

(2). What criteria and factors have been used by automated methods for their operational definitions?

Each of the existing algorithms and methods of feature extraction carries a definition. By analyzing these algorithms, this research decodes the buried or hidden operational definitions. Dissection of these operational definitions may reveal what criteria have been used and what is missing.

(3). What are the performance levels of these approaches?

A series of tests are proposed to evaluate the rationale and performance of these approaches. The resulting ridge and valley lines will be analyzed and compared.

(4). Which approach performs most closely to manual extraction?

By examining the results of each different approach, this research attempts to identify those approaches that more closely reflect human conceptual models of ridge and valley lines. The identified approaches should be useful to general purposes of terrain features extraction.

(5). How and why do these approaches perform differently?

The different performance between approaches can be attributed to many sources such as human cognition, nature of terrain model, and implementation. The identification of such sources will facilitate our understanding on feature extraction and enhance the development of automatic extraction methods.

The purpose and influence of this research is multi-faceted. Immediately, findings from this research will provide a more thorough solution and understanding of automatic extraction of terrain features. In a broader sense, this research deals with a wide range of subjects relevant to terrain feature extraction, a geographic task. It covers topics on geographic data, human factors, and implementation in computers. Through the process, the many problems regarding computers and geographic applications will surface. This will result in a better understanding of what computers can do and can not do, therefore promoting a better use of computers in geographic analysis. Theoretically, this research deals with human cognition and spatial reasoning. There are attempts in developing automatic image understanding system (Erickson *et al.*, 1984; Goodenough *et al.*, 1987; McKeown, 1987). Such tasks are related to human cognition and conceptual models. Findings of this current research on conceptual and cognitive aspects of spatial reasoning may be fruitful to these research interests.

Chapter 4. Conceptualization

4.1 Feature Extraction and Categorization

Every automatic method for the extraction of terrain features requires clear and implementable definitions of the features. For example, a computer needs to be programmed to what exactly a 'ridge' is before it can extract ridges. Thus, the first task in the extraction of terrain features is to define the features. Although there exists formal descriptions of terrain features in many dictionaries, these descriptions are often ambiguous or circular and hence unsuitable to be implemented in computers. For example, one definition of a ridge in the *Webster's Third New International Dictionary* (Gove, 1976) is: an extended elevation between valleys. One question arises from this definition: what is a *valley*? Without knowing what constitutes a valley, this definition is meaningless. Another definition in the same dictionary specifies that a ridge is 'a top or upper part especially when long and narrow'. But then, how long is long and how narrow is narrow? In regards to definition, the challenge facing the automatic extraction of terrain features is twofold: first, to clarify the characters of terrain features; second, to adapt these characters, mostly expressed verbally, into operational digital definitions that can be implemented in computers.

This chapter first investigates the properties of terrain features. Through references to dictionary descriptions and interviews with experts, the characteristics of terrain features are explored. The analysis of 'meaning' and 'definition' is part of a categorization process that delineates boundaries of categories. 'Ridge' and 'valley' are categories that compose certain parts of the earth's surface and possess particular characteristics.

Categorization has been a topic of interest in many academic fields such as cognitive science, linguistics, and philosophy. Theories of categorization should be useful for a better understanding on the definition of terrain features. The characteristics of terrain features identified from dictionaries and interviews will be analyzed with reference of these categorization theories. The analysis will address the problems in defining features and suggest possible solutions for the development of operational definitions to be used in computers.

4.2 Characters of Terrain Features

In the literature of automatic extraction, the terminology used to describe terrain features is diverse. The term 'ridge' has been commonly used in the literature of terrain feature extraction (Douglas, 1986; Riazanoff *et al.*, 1988), although the use of 'divide' can also be found in literatures (Band, 1986; Fairfield *et al.*, 1991) For the counterpart of ridges, many different terms have been used, such as 'valley' (Riazanoff *et al.*, 1988), 'channel' (O'Callaghan *et al.*, 1984; Douglas, 1986), and 'drainage network' (O'Callaghan *et al.*, 1984; Fairfield *et al.*, 1991). The definitions of these terms are to be explored, in this context, with reference to several dictionaries.

Besides the aforementioned terms, many other related terms are found in dictionaries. For example, the term 'valley' is used to define many other terms, such as 'ravine', 'gully', and 'gulch'. The term 'ridge' is used to explain 'divide', which is related to 'water parting'. To obtain a complete picture of these terms, a review on the descriptions of relevant terms would be helpful. The definitions for *ridge* and associated terms are quoted below.

ridge:

a) The line of intersection at the top between the opposite slopes... (Gove, 1976, p.1953).

b) A range of hills or mountains or the upper part of such a range(Gove, 1976, p.1953).

c) An extended elevation between valleys (Gove, 1976, p.1953).

d) A top or upper part especially when long and narrow (Gove, 1976, p.1953).

e) A long narrow upland, with steep sides... (Monkhouse, 1970, p.297).

f) Ridge is applied loosely to any long narrow rise in the ground, hills etc. (Stamp, 1966, p.353).

divide:

a) A dividing ridge or section of high ground between two basins or areas of drainage (Gove, 1976, p.663).

b) A ridge or area of high ground between river basins (Monkhouse, 1970, p.110).

water parting:

a) A summit or boundary line separating the drainage districts of two streams or coasts (Gove, 1976, p.2584).

b) The boundary line between two or more drainage basins (Stamp, 1966, p.458).

Based on these descriptions, a ridge can refer to a linear feature as well as an area feature. It is a line when considered as the intersection of opposite slopes or a divide between water basins. It is an area feature when considered as 'a top of upper part' or 'an extended elevation between valleys'. The definitions of 'valley' and its related terms follows.

valley:

a) An elongate depression of the earth's surface and commonly situated between ranges of hills or mountains and often comprising a drainage area (Gove, 1976, p.2530).

b) An elongated depression sloping towards the sea or an inland drainage basin, usually though not always occupied a river (Monkhouse, 1970, p.365).

c) An elongated depression, usually with an outlet, between ranges of hills (Stamp, 1966, p.445).

gully:

a) A miniature valley or gorge worn in the earth originally by running water through which water usually runs only after rains (Gove, 1976, p.1011).

b) A small ravine in the face of a precipice (Gove, 1976, p.1011).

c) A well-defined waterworn channel on a hill-side (Monkhouse, 1970, p.167).

ravine:

a) A small narrow steep-sided valley that is larger than a gully and smaller than a canyon and is usually worn down by running water (Gove, 1976, p.1887).

b) A narrow steep-sided valley, but larger than a gully or a cleft (Monkhouse, 1970, p.288).

gorge:

a) A narrow steep-walled canyon or a particular narrow steep-walled part of a canyon (Gove, 1976, p.980).

b) A ravine with steep rock walls (Gove, 1976, p.980).

c) A deep, steep-sided, rocky river valley (Monkhouse, 1970, p.160).

canyon:

a) A deep narrow valley with precipitous sides characteristic of regions where downward cutting of the streams greatly exceeds weathering (Gove, 1976, p.329).

b) A deep, steep-sided gorge with a river at the bottom; manly found in arid or semi-arid areas, where a rapidly eroding river maintains its volume from snow-melt on distant mountains....(Monkhouse, 1970, p.56).

stream:

a) A body of running water flowing in a channel on the surface of the ground... (Gove, 1976, p.2258).

b) A body of flowing water, covering all scales from a small rill to a large river. Hence the term is used to denote all the characters of processes and landform resulting from streams (Monkhouse 1970, p.334).

c) Any river or course of running water of current in the sea, as Gulf Stream (Stamp, 1966, p.399).

channel:

The deepest part of a river-bed, containing its main current, naturally shaped by the force of water flowing in it (Monkhouse 1970, p.64).

thalweg:

A 'valley-way', but used to denote the longitudinal profile of a river. (Monkhouse, 1970, p.347)

watershed:

a) A region or area bounded peripherally by water parting and draining ultimately to a particular water course or body of water (Gove, 1976, p.2584).

b) The catchment area or drainage basin from which the waters of a stream or stream system are drawn (Gove, 1976, p.2584).

c) The line separating headstreams which flow to different river systems...; in the U.S.A.,this is equivalent to a divide (Monkhouse, 1970, p.370).

A valley is a basic terrain feature that has many derivatives in various shapes and scales, such as: ravine, gully, canyon, and gorge. For this study, the term 'valley' will refer to all the various derivatives (shapes and sizes) of a valley. Apparently, the term 'valley' has many connotations. It is part of a water basin and embraced by its surrounding ridges. Although a valley normally is considered as an area feature, it is closely associated with linear terrain features such as streams, channels, or thalwegs. This linear association is even stronger when a valley is narrow and deep. In a mountainous area, where most analyses of terrain are being conducted, the distribution of valleys also suggests the presence of stream networks. A valley line which indicates the location of a valley is in fact a thalweg of that valley. In this sense, the use of 'stream', 'channel', or 'thalweg' is closely related to 'valley'. Yet, the terms 'stream' and 'channel' strongly suggest the presence of water which may not be a major concern to many other applications. The term 'valley line' is more conclusive in reflecting the characteristics of this group of terrain features.

The dictionary definitions primarily describe the spatial location and geometry of terrain features. Not all of these descriptions are useful. Those circular definitions that use valley and ridge to define each other are not very useful except to show the duality of these two features. Some of the definitions related to geometry are too vague to be implemented in computers. To further investigate the properties of terrain features, interviews with specialists from various fields were conducted. By knowing how and why researchers use terrain features in their domain of study, the properties of these features can be better perceived. The following section documents the findings of the interviews.

4.3 Specialists Interviews

Six specialists with experience in soil science, geomorphology, forest management, hydrology, and cartography, were asked: 1) why they need to delineate terrain features, and 2) how they use terrain features in their research. The answers should reveal the purposes of using terrain features in research and will help to determine their definitions. The results of the interviews are summarized below.

Two interviewees in the field of engineering geomorphology indicated that ridge and valley lines are basic components of their analyses of slope stability. Ridge and valley lines are the boundaries of different slope complexes. By delineating ridge and valley lines, they can roughly delineate slope complexes of various aspects. These units are further divided into sub-units based on the density of contours. From this delineation, slope and aspect maps can be produced. The slope and aspect map of a study area then can be integrated with the dip and strike of geological structures. By analyzing these data, geomorphologists can identify areas with a high potential of landslides. In their view, a ridge line is the intersection of opposite slope complexes while a valley line is the intersection of two facing slope complexes. Such views reveal the important role of geometric properties in defining ridges and valley lines.

In soil science, the development of soils is closed related to landform type, and therefore landform patterns and topography are of general concern. In particular, soil erosion is related to slope and slope length. The Universal Soil Loss Equation, a widely used equation to predict soil erosion, is expressed as (Gerrard, 1981, p.35):

$A = RKLSCP$ where A = average annual soil loss;
R = rainfall factor;
K = soil erodibility factor;
L = slope length factor;
S = slope steepness factor;
C = cropping and management factor;
P = conservation practice.

The slope length is the distance from a ridge line to either the slope foot or valley bed; therefore, the location of ridge lines is needed for the calculation of slope length.

In the field of water resource management, the boundary of a water basin is important, for it indicates how much overland flow will run into a drainage system. This boundary combined with drainage patterns and stream ordering is used by hydrologists to predict the pattern of storm hydrography and floods. From an environmental perspective, the boundary of a water basin is important for the control of water quality. With the delineated watershed, environmental managers can identify a potential source of water pollution and take appropriate measures.

The location of ridge and valley lines is a factor in timber management as well. One interviewee indicated that in many cases logging is not allowed within certain distances from ridge tops and valley beds. Foresters, in order to comply with natural resource legislation or regulation, need to plot a buffer zone around ridge and valley lines for the management of forestry.

For cartographic illustration, ridge and valley lines are important in portraying the terrain. On contour maps, the plot of ridge and valley lines will enhance the readability of maps by revealing the general structure of an area. Another widely-used representation of terrain is a shadow relief map. By mimicking the diffuse reflection of sunshine on the ground, this map helps a reader to comprehend the pattern of relief of the ground surface. Ridge lines are the natural boundary of different reflection zones and thus are the boundaries of different tones on the shadow-relief map. Cartographers often start shading along ridge lines. A pre-process to delineate ridge lines is often conducted. In the production of topographic maps or relief maps, the delineation of ridge and valley lines is a fundamental step.

These interviews provided multi-perspective descriptions of ridges and valley lines. This author did not explicitly ask for definitions during interviews based on an assumption that general descriptions and applications of terrain features can provide a broader picture of terrain features. These various applications and general descriptions collected from interviews will be summarized in next section. It is, however, questionable whether these properties are complete and precise enough to induce suitable operational definitions for the implementation in computers. This author will analyze these properties with reference to categorization theories.

4.4 Analysis of Definitions of Terrain Features

Three major properties of ridges and valleys emerge from interviews and the descriptions in dictionaries. They include hydrologic, spatial, and geometric perspectives. From a hydrological perspective, a ridge is the divide of overland flow and a valley line is where overland flow accumulates. In a spatial context, a ridge is a line along the top of mountain and a valley line is along the lowest part of a valley. From a geometric perspective, a ridge and a valley line delineate the boundaries of opposite slope units. Deduced from this geometric perspective, a ridge and a valley can be defined as a linear feature with a convex downward shape along its cross profiles. In conclusion, four definitions of 'ridge' and 'valley' have been loosely identified. The variation of these four definitions leads to an analysis of categorization.

In the traditional view of categorization, a category is composed of entities that share common properties, which are necessary and sufficient criteria to define a category. Each category is thought of as having a clear and fixed boundary, which may be a set of well-defined criteria (Lakoff, 1987, p.16). Many questions will arise if we apply this traditional theory to the examination of the four definitions of ridge and valley. Are they necessary and sufficient to define ridge or valley line? Is it possible to delineate a clear and fixed boundary separately for ridges, or valley lines? Are these properties consistent regardless of the variation of terrain types and area size? The answers to these questions are probably negative. Recent advances in cognitive science and experimental psychology develop many new theories on categorization. These theories include prototype effect of category (Rosch, 1975, 1978), *ad hoc* categorization (Barsalou, 1983; Lakoff, 1987, p.45), and fuzzy categorization (Lakoff, 1987, p.21). Within these theories, as explained below, the fuzziness of categorization is the most relevant concept to the categorization of terrain features.

Much ambiguity of the definition of ridges/valleys can be contributed to the fuzzy nature of categorization. Fuzzy set theory was first introduced by Zadeh (1965) and has been widely applied to various fields since. Instead of being precise and clear, many human concepts are fuzzy, *e.g.*, tall vs. short, steep vs. gentle. In classical theory, an element is either a member or not a member of a set; in fuzzy theory, membership is not a binary yes-or-no phenomenon, but a graded property. A six-foot-high man is more likely to be considered tall compared to a five-foot-eight man, yet it is difficult to define how tall is tall. The effect of fuzziness to human categorization can be illustrated by the experiment conducted by Labov (1973). Labov drew many similar objects with varying heights and widths to show a graded variation between cup and bowl, as in Figure 4.1.

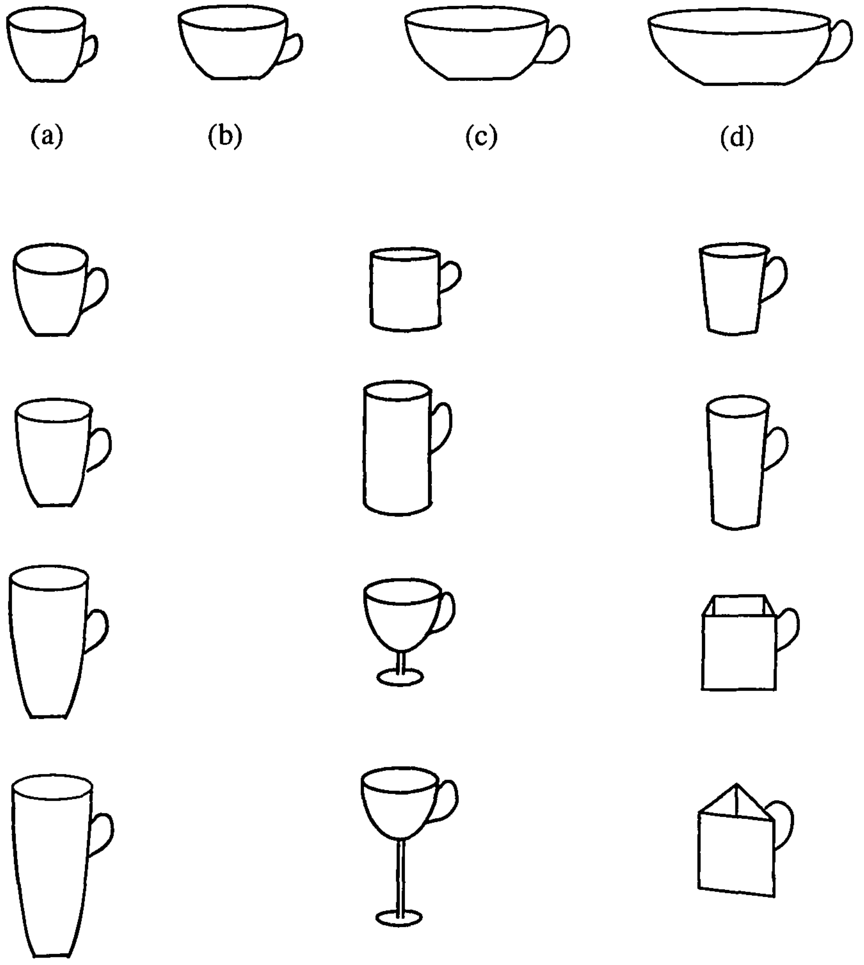


Figure 4.1: Cups of various shapes (redrawn after Labov, 1973).

Labov asked his experimental subjects to categorize the four objects at the top row of Figure 4.1 as cup or bowl. These four objects have the same height while their ratios of height to width vary. He asked the subjects of this experiment to decide if these objects were cups or bowls (1) when the objects are neutral, i.e., the objects are empty, and (2) when there are food inside these objects. The result of categorization showed that there was no clear boundary between the categories of bowl and cup.

Some subjects thought item 2 was a cup while others thought it as a bowl. Also, the boundary is not fixed. Many subjects changed the categorization of bowls and cup when the contents of these objects changed. This experiment shows the fuzziness of boundaries between categories and an *ad hoc* effect on categorization.

The vagueness of categorization identified by Labov's experiment pertains to the definition of terrain features as well. The ground surface is composed of slope units with continuous variation, from a very steep cliff to gentle horizontal plain, and from huge mountains to trivial reliefs on the ground. Such variations are presented, by mimicking the work of Labov, in the sample plot of Figure 4.2.

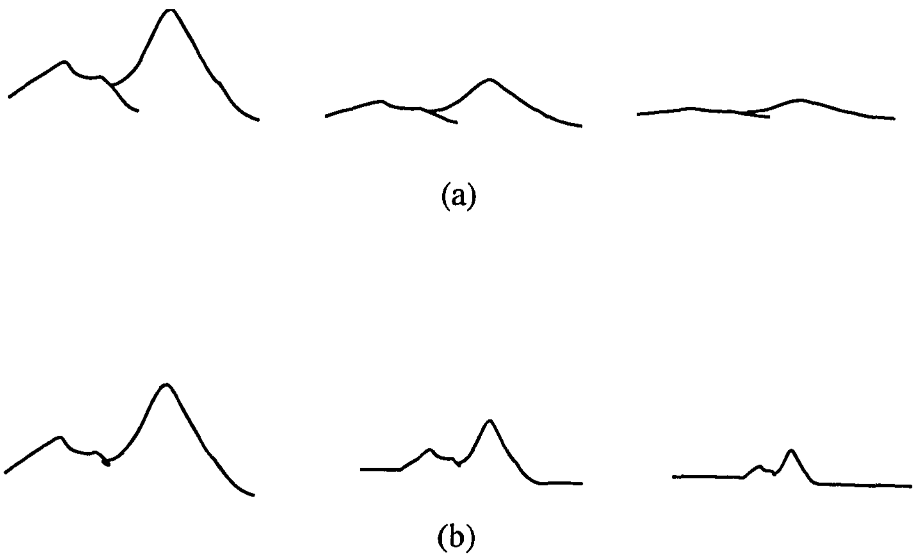


Figure 4.2: (a) Cross profiles of various shapes; (b) cross profiles of various size.

The three profiles in Figure 4.2a show a continuous variation of steepness in the shape of terrain profiles. The profiles in Figure 4.2b show a variation in the size of terrain objects. Shape and size are two apparent factors in defining terrain features. It is, however, improbable and implausible to derive a set of consistent measurements that can separate the steep from non-steep for universal applications. Even if cross profiles are in a same shape, the magnitude of these profiles affects human interpretation of them. The bigger ones will be more likely to be interpreted as ridges compared to smaller ones. Besides shape and size, in real-world situations, the contents of an area or its surrounding landscape will affect people's interpretation of terrain features. A peak of 50 meters may not be significant in the Rocky Mountains, yet it could be a significant feature if located an American Midwest plain state. Such context-oriented effects highlight the *ad hoc* effect in the definition of terrain features and makes the delineation even more complicated. Given such vagueness and uncertainty in extraction, human interpreters have no problem, with some subjectiveness and arbitrariness, in extracting the desired features for their application. How to deal with the various requirement of difference application is a challenge facing the development of automatic extraction method.

4.5 Conclusion

This chapter explored the definitions of ridges and valley lines through the descriptions in dictionaries and by interviews with domain experts. It is found that the meanings of ridges/valleys are multi-faceted. *Ridge* and *valley* have different meanings to practitioners with different purposes of extraction. In a cognitive term, these variations may reflect an *ad hoc* categorization of terrain features. In another theory of cognitive science, characteristics that used to describe ridge and valley, *e.g.*, shape and size, are often graded and fuzzy and unlikely to serve as a precise measure. To be useful to the various applications, a satisfactory method of automatic extraction should be able to accommodate into its implementation these *ad hoc* purposes and flexibility in size and shape. The manual extraction of ridge/valley lines has been an subjective and *ad hoc* process and should remain so in automatic extraction. The type and amount of extracted features should be left as an option to the individual user, based on his/her purpose of extraction, of automatic methods.

Four groups of automatic methods have been identified in Chapter 3. Each of these groups of methods carry a unique operational definition of ridge and valley lines. The interviews and reviews in this chapter identified several characteristics pertaining to the definition of ridge and valley lines as well. The next four chapters investigate each group of methods separately. The characteristics embedded in each operational definition of ridge/valley lines will be examined. The performance of each method will be tested through implementation on computers with a sample data set. A series of analyses and experiments will reveal how the difference of definition affects the results of extraction. The performance of each method is subjected to further comparison and analyses.

Chapter 5. Symbolic Approach

5.1 Introduction

This chapter examines the feasibility of a symbolic approach for the automatic extraction of ridge and valley lines, an approach first proposed by Palmer (1984) and followed by Frank *et al.* (1986). The term *symbolic* was adopted to denote that the approach explicitly expressed the relations between different terrain features, and used them to define and extract terrain features automatically. The development of such approach was made possible largely attributed to the availability of expressive computer language such as Prolog, which makes explicit representation of rules and relations much easier compared to other computer languages.

Palmer (1984) adopted the *symbolic approach* to delineate several terrain features: peaks, pits, ridges, and streams (termed as valley lines by this author). The program was written in Prolog and the terrain data was represented in a TIN structure. It was successful in the sense that it delineated all the designated terrain features in his simple data set, however the data was well below the complexity of real world terrain. As mentioned in Chapter 3, an automatic method for the extraction of terrain features needs to integrate the nature of terrain models into the algorithm development. A method that can handle a simple data set may fail on a complex data set. Whether this method is feasible for real world applications will be analyzed and tested.

This chapter first examines the Prolog code to determine the operational definition of terrain features embedded in the program to determine whether the operational definition reflects a specific human conceptual model. The nature of a terrain model is an indispensable factor to the operational definition of terrain features. As the program was developed to work on a TIN structure of a DTM, a further investigation of TIN is necessary. Through a cross reference of TIN and this program, a final evaluation of this method will conclude this chapter.

5.2 Analysis of Operational Definition

The terrain data and program in Palmer's paper was written in the Prolog language which is a computer language suitable for dealing with objects and relations. A Prolog program is composed of 'facts' and 'rules'. The terrain data will be 'facts' in this case. In a TIN structure, terrain data can be composed of nodes, segments, and triangles, or cells in Palmer's terms. For a TIN in Figure 5.1, Palmer specified these items in Prolog. Some sample facts are listed below.

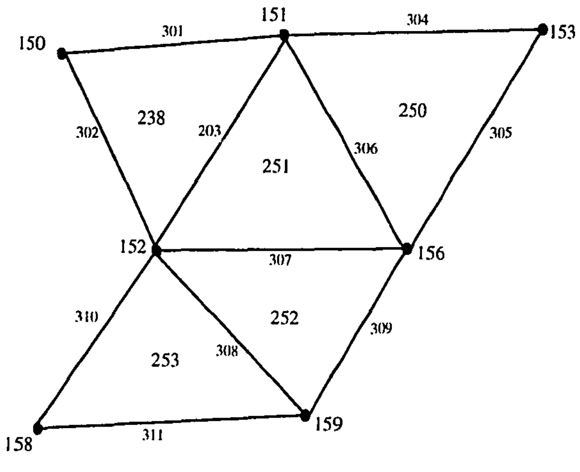


Figure 5.1: Part of a sample TIN matrix.

```
node ( 150, 6987, 2152, 1010, [240,238],260).
node ( 151, 8482, 1861, 1240, [250,238,240]).
segment (301, [150,151], [238,240]).
segment (302, [150,152], [260,238]).
cells (238, [150,151,152], [240,250,260], [301,303,302], [1,-.3, -.7]).
cells (250, [151,153,156], [270,280,238], [304,305,303], [-.7,-.3,1]).
```

Each node fact indicates the node's identifier (ID), x, y, and z coordinates, and the cell(s) it belongs to. Each segment fact specifies the segment's ID, the Ids of the two nodes at the two ends, and the Ids of the cells (triangles) on both sides of the segment. For a cell, the fact specified its ID, the Ids of the three nodes on the corners, the Ids of three adjacent triangles, the Ids of the segment on three sides and the flow direction along each segment. These facts are very informative, including coordinates and topology of the TIN matrix.

Built upon these facts, Palmer defined 'ridges' and 'streams'. A ridge is defined as:

```
cell_flow ( C, S, F) :- cell ( C, _, [S, _, _] ,_, [F, _, _]).
cell_flow ( C, S, F) :- cell ( C, _, [_, S, _] ,_, [_, F, _]).
cell_flow ( C, S, F) :- cell ( C, _, [_, _, S] ,_, [_, _, F]).

ridge(S) :- segment (S,  _, [C1,C2]),
            cell_flow (C1, S, F1), F1 >0,!
            cell_flow (C2, S, F2), F2 >0.
```

To understand this definition better, an introduction to the format of Prolog language will help. Prolog rules are composed of a head and a body, separated by a colon and dash (:-). The head can be taken as goal to be realized, based on the body at the right hand side. The body can be composed of logic expressions or sub-goals which need to be subsequently evaluated. Capital letters in a rule symbolize variables and an underscore (_) means any value. The rules of 'cell_flow' are to determine whether water would flow in or out of a cell across each of the segments of a cell. Since a segment can be the first, second, or the third item on the segment list of a cell, it takes three rules to specify the 'cell_flow'. The rules for a 'ridge' specify that for an item S to be a ridge, it must be a segment and the flow direction of this segment to its two adjacent cells must be positive, which can be determined the elevations of four adjacent nodes.

By querying a data base of facts with these rules, the user can test whether another fact is true or not. For example, to check whether segment 303 is a ridge, the user can key in:

ridge (303).

The Prolog interpreter will work backwards, starting from a goal to evaluate the body at the right, while the body itself may be goal itself and need more evaluations. To evaluate 'ridge (303)', the interpreter will evaluate 'segment (303, _, C1,C2)' first, which is validated since segment 303 has been defined in the fact in previous part of program. The evaluation of segment also associates the two cells, 238 and 251, with segment 303 which are part of the fact. After identifying these two cells, the interpreter will find the flow directions of segment 303 to cell 238 and to cell 251. If both directions are positive, which means waterflow would run from the segment into the cells, then the segment is a ridge. The query can be used to find all the segments that fulfill the definition of ridge by inquiry with a variable, such as:

ridge (A).

The interpreter will list all the ridges defined in these way through pattern matching.

A further analysis of this operational definition reveals that this method defines valley and ridge lines based on the shape of the cross profile. If water flows from the segment to its adjacent triangles, then it is a ridge. If the segment is the highest edge of the three sides it is assumed water will flow from that segment back into the cell. If a segment is higher than both its adjacent nodes, the two cells will pose an 'A' shape in a cross profile. If the flow directions are from the cells to the segment, then the segment will be a valley line and there will be a 'V' shape profile instead. To express this in another way, this definition is based on comparisons between the segment and its two adjacent nodes, the nodes opposite of the segment. If both adjacent nodes are lower than the segment, then the segment is a ridge. As long as the two adjacent nodes are lower, it does not matter how low or how high, it will be classified as a ridge.

This definition does not specify scale and steepness. As discussed in previous chapters, scale is a major factor in the representation of a terrain feature. A TIN can represent a terrain surface with varying scale and detail. The definition and criteria in this method are rather primitive; they rely on TINs to deal with the generalization and scale problem of terrain features. To apply this method in real world terrain, the ability to create a TIN that adequately represents ground surface is critical. To pursue Palmer's method of feature extraction requires an investigation on the creation and nature of a TIN structure.

Currently, a TIN structure of DTM is not generally available. DEMs of regular-grid format are almost the sole source to create TINs. The following section will investigate this creation process and evaluate the reliability of TIN.

5.3 Constructing a Triangulated Irregular Network

The extraction techniques of TINs is a topic of interest to many researchers. In general, the generation of a TIN from regular-grid DEMs involves two steps: first, to select the nodes as basic elements of networks; secondly, connecting these nodes into triangles. The following section will investigate the node selection process.

5.3.1 Selection of Nodes

Two methods to select nodes are widely recognized: the very important points (VIP) method and hierarchical method. The VIP method selects nodes based on the a local relief. For any point inside the DEM data, a 3 by 3 window of a DEM is taken for analysis. The average elevation of the eight neighbors is computed. This average is compared to the elevation of the center point. A big difference between these two values suggests a major change with trend of relief occurs at the central point, such as on the ridge top, valley bed, or slope break. Although a large difference between the values could also result from an isolated relief feature. There will usually be some difference between the elevation of the average and that of the center point and users may specify the number of nodes to be selected based on the ranking of deviation. In this way, the user can control the size of extracted TIN. The selection may also be based on a threshold value and all those points with a deviation exceeding the threshold value may be chosen as nodes.

The second widely accepted method to choose nodes is the hierarchical method. The method is a 3-dimensional sibling of the Douglas-Peucker algorithm (Douglas and Peucker, 1973) for line generalization. The Douglas-Peucker algorithm begins the line generalization by first connecting the two ends of a line to be generalized as a straight segment. The distance between each selected point on the original line to the straight segment is calculated and compared. The farthest point from the segment is chosen as an intermediate end point and connected to the two ends of the original segment. The segment is therefore split into two segments, as Figure 5.2 shows.

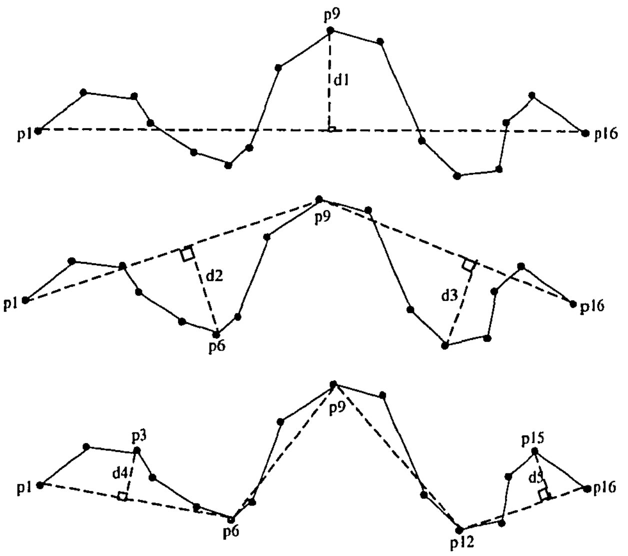


Figure 5.2: The Douglas-Peucker algorithm for line generalization.

Within each of these new segments, the same procedure is taken to identify the farthest point and break the segment into two further divisions. This procedure is recursively taken until the distance between the farthest point and the correspondent segment is within a pre-defined tolerance value. Derived from this generalization method, a node selection procedure for three dimensional planes is being used for the TIN construction (Lee, 1989, 1991).

When applied to 3-D data, the rectangle area of a regular-grid DEM is arbitrarily divided into two triangles, by connecting two diagonal corners into a segment. Within each triangle, the distance between each point of the original DEM and the plane defined by the three nodes is computed. The point most distant from the plane is chosen as a new node and the original triangle is split into three triangles. Within each resultant triangle, the computation of distance and selection of new nodes is recursively performed, until the distance between the farthest point and the triangle is less than the pre-set tolerance value.

5.3.2 Connection of Nodes

After selection, these nodes need to be connected by segments and formed into triangles. The connection will affect the representation of the terrain. Most methods will first sort the nodes based on their x,y coordinates and then connect them. The Delaunay triangulation is a widely used method to connect nodes to form facets. The It is based on a circular criterion that the circumcircle through the three nodes of any triangle does not contain any other nodes. If the circumcircle of a triangle contains other nodes, the triangle will be re-configured to fulfill the circular criterion, as shown in Figure 5.3.

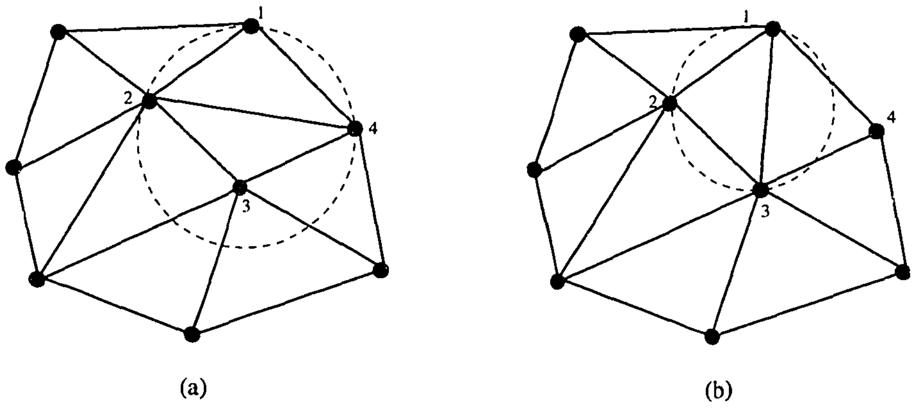


Figure 5.3: The circular test of Delaunay triangulation.

A merit of Delaunay triangulation is that it will maximize the minimum interior angle of the triangle formed, thus nodes can be linked to their nearby neighbors and form equiangular triangles (Lee *et al.*, 1980; De Floriani *et al.*, 1984; Lee, 1991).

To those nodes selected by the VIP method, there is no underlying relations between them. The adoption of the Delaunay triangulation to connect those nodes may be a reasonable solution, compared to no triangulation criterion at all. On the other hand, the nodes selected from the hierarchical method are associated with the triangle, and the three nodes of the triangle, from which they are selected. Whether we can skip these embedded triangles and adopt the same criteria as Delaunay test is questionable and requires further study.

5.4 Tests

To compare the variation of these methods, a series of tests were conducted. A 1:24,000 DEM of Kaneohe, Oahu produced by USGS was used to build several TINs. The Kaneohe DEM contains of 256 by 256 pixels, or 65,356 points. For the first test, the VIP method was used to choose 2000 nodes and 5000 nodes separately. These nodes are connected into TINs with a Delaunay triangulation. The TINs are shown in Figure 5.4 and 5.5.

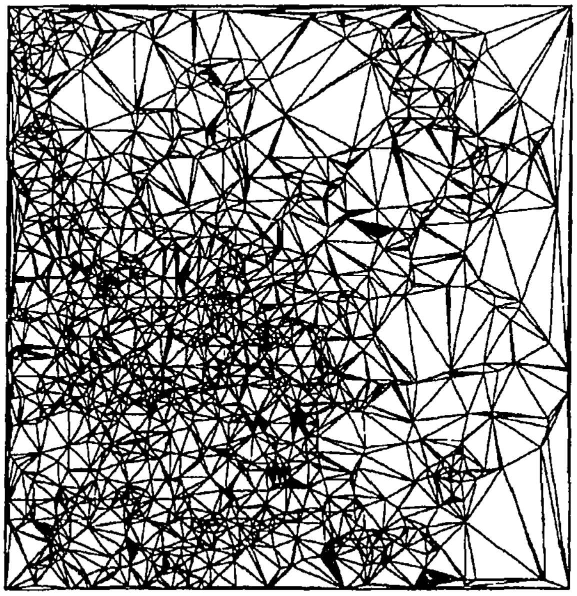


Figure 5.4: A TIN matrix of 2000 nodes generated by VIP method.

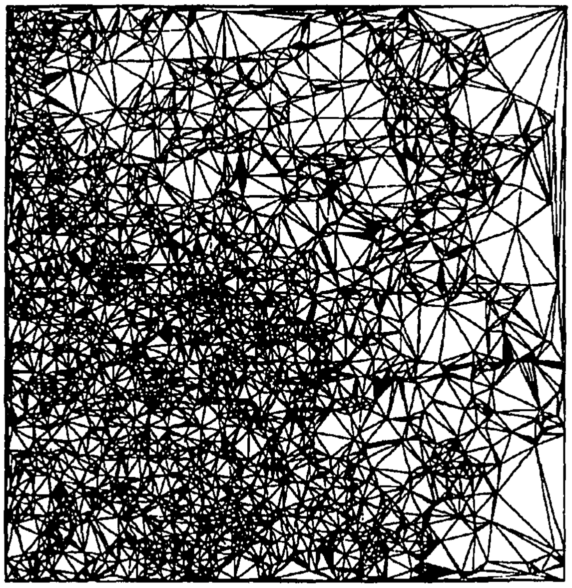


Figure 5.5: A TIN matrix of 5000 nodes generated by VIP method.

The same data set was used to select nodes based on the hierarchical method. Two different threshold values of 50 and 80 meters were used for two different tests. These two tests selected 1578 and 3548 of nodes respectively. The Delaunay triangulation was used to connect these two sets of nodes, as illustrated in Figures 5.6 and 5.7.

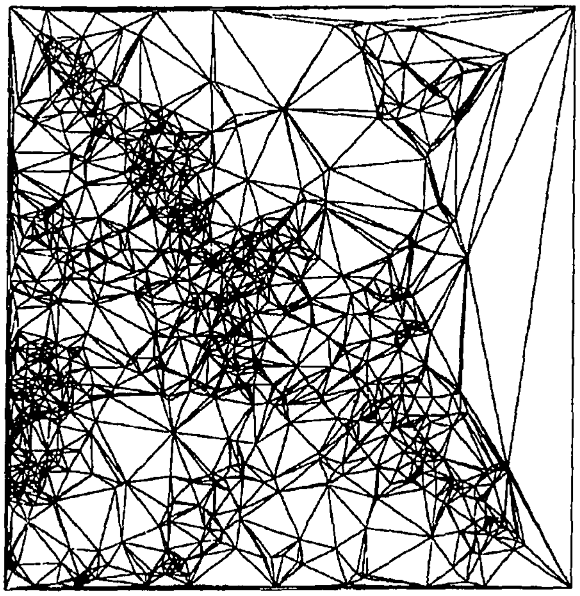


Figure 5.6: A TIN matrix with nodes selected by hierarchical method using a 50 m threshold and connected by Delaunay triangulation.

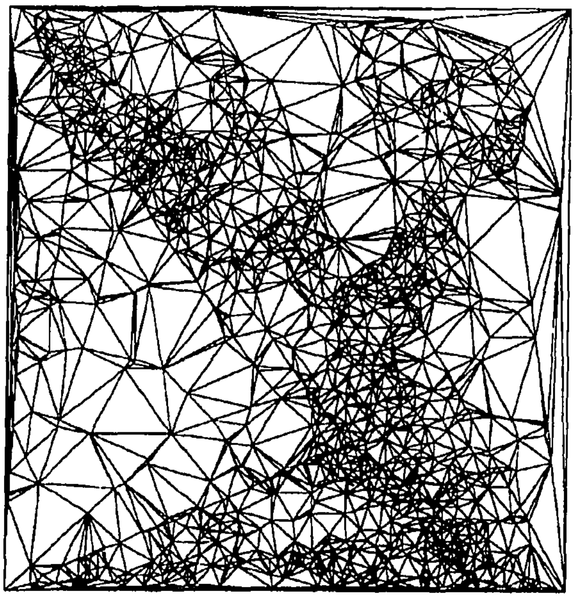


Figure 5.7: A similar TIN of Figure 5.6 except using a 80 m threshold.

The same sets of nodes from the hierarchical methods were connected based on the original triangles that each nodes were chosen from, the result is shown as Figure 5.7 and 5.8 at below.

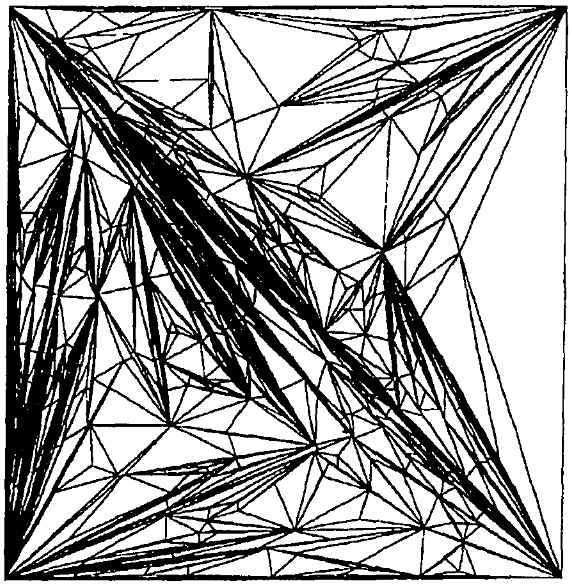


Figure 5.8: Same nodes of Figure 5.6 connected by their original triangles.

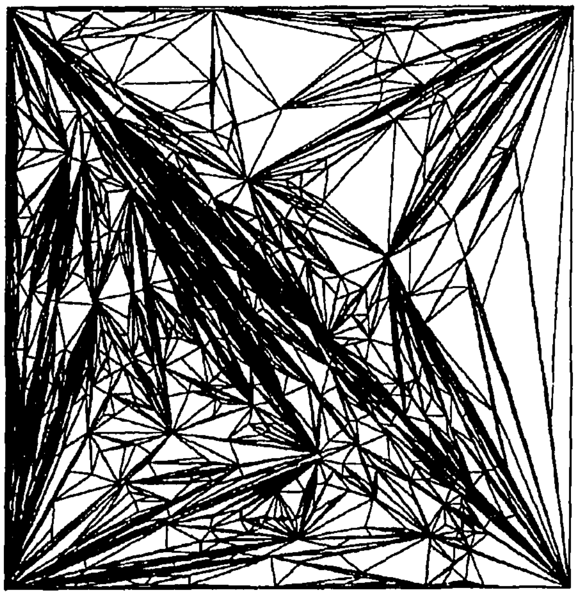


Figure 5.9: Same nodes of Figure 5.7 connected by their original triangles.

There are various concerns relating to the preservation of terrain characters, such as elevation and ground surface pattern. The resultant TINs from these experiments are very different. It is questionable whether they resemble the original surface or preserve the elevation. Following section investigates the representativeness of these various TINs.

5.5 Discussion

The VIP method does not guarantee a fit between resultant TINs and the original DEM. The nodes selected from the VIP methods are important in a local area of 3 by 3 window, yet this does not imply they will be critical to global representation of the DEM. On the other hand, a major relief of a large area may contain no VIP at local areas. Take the profile in Figure 5.10 as an example.

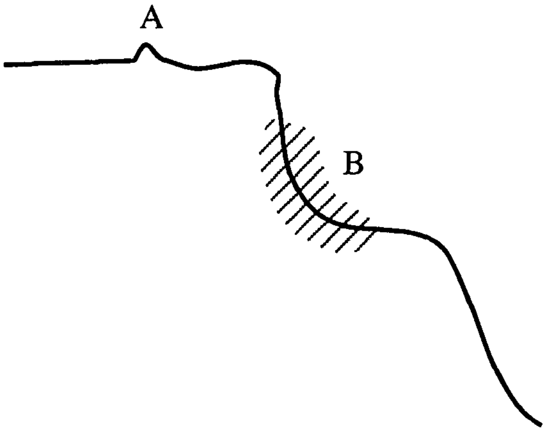


Figure 5.10: A sample cross profile.

Point B is on a major curve of smooth transition. Yet because this transition is smooth, we cannot identify any significant difference between the point and its neighbors, therefore no VIP will be detected. On the other hand point A can be selected as a VIP because it fits the requirement of a VIP although it is not as important as point B in representing the whole surface. The VIP method fails to see these important differences.

The connection of nodes into triangles critically affects the resultant terrain surface in terms of elevation, slope, and aspect. The four points in Figure 5.11 may represent a ridge, or a valley, depending on how we connect them.

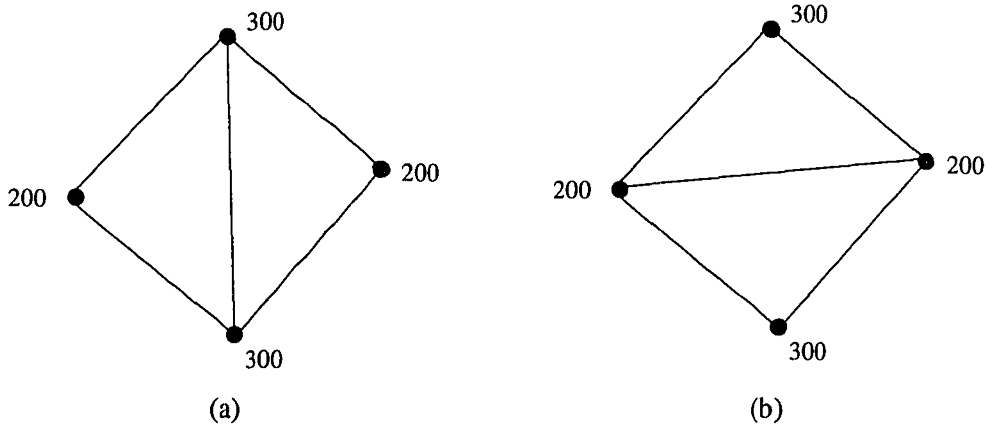


Figure 5.11: Different connections of TINs with same nodes.

The connection itself decides the final configuration. The correct connection should be based on the ground relief. The Delaunay network on VIP points is not based on actual topography, but on horizontal angles between points. There may be merit in connecting nodes with nearby neighbors and to building acute triangles. Yet, the sole reliance on geometric pattern for triangulation is insufficient for the TINs to represent ground surface. This theoretical deficiency of automatic TIN generation is yet to be resolved.

The hierarchical process selects nodes based on the distance between points to their embracing triangle. The resulting triangles should preserve the elevation within an error margin. Yet, this guarantee can only be held when nodes are connected based on the original triangles they were chosen from. The selected nodes are important in representing the original surface because they are farthest away from the particular triangle. Without the embracing triangle, they may not be important at all. Take the line in Figure 5.12 as an example to illustrate this problem. Point A is selected because it is the farthest point from the segment CD. If we chose A and then did not plot CD, then the representativeness of A in such segment is questionable.

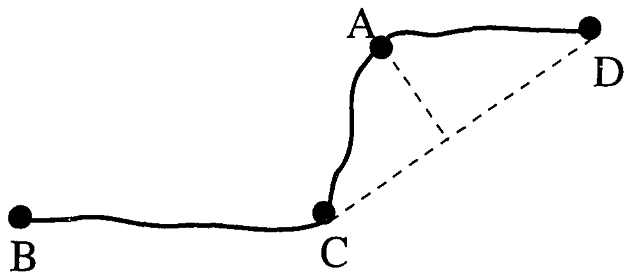


Figure 5.12: The node selection process in a segment.

Theoretically, those points selected by the Hierarchical method should be linked with their original triangles. The TINs in the Figure 5.8 and 5.9 resulted from such connection. They look very awkward and unnatural with many long and obtuse triangles. This representation may preserve the original elevation to a tolerance level, yet it distorts the terrain pattern. This method recursively breaks triangles into 3 smaller one's with a new nodes, either lower or higher than the triangle. This new configuration will be a pyramid by the nature of picking the farthest point. Consequently, the aspect of these three triangles will be facing away from each other and the result will be unnatural and unfaithful to the original terrain surface. On the other hand, if we use the Delaunay triangulation to connect these nodes, then the role of these nodes will be of question first of all, let alone whether the connection will make sense or not.

By the nature of TIN, adjacent triangles tend to often differ in aspect and slope. Some of these differences may be reflecting variation on the ground surface, yet not uncommonly, some of these differences are introduced by the generation of the TIN. The latter case makes the TIN very unnatural and unfaithful in representing ground patterns. In a recent attempt to derive slope lines from digital terrain data of vector format, Chou (1992) also found that the TIN he created was unnatural in representing the earth's surface.

From theoretical analysis, several problems of node selection and connection have been identified. From implementation and testing we have seen wide disagreement in the representation of TINs. From the graphic illustrations we also see the failure of TINs in rendering the original surface. All these show a revision on the TIN construction method is needed. Ideally, each facet of TIN represents a homogeneous slope unit on the ground. To represent the ground faithfully, the TIN should be built upon and reflect the true boundaries of slope units on the ground. To fulfill this goal, a TIN should be built upon lines of ridges, drainage networks, and slope breaks (Douglas 1986). These lines delineate the slope units and should serve as the backbone of the TIN structure. In other word, the locations of valleys, ridges, and slope breaks are prerequisite data for a faithful representation of terrain in TINs.

5.6 Conclusion

The computer program by Palmer in fact compares the elevations of one segment with the elevations of its two adjacent nodes, although this comparison is not explicit in the program. The use of Prolog with a TIN is the major factor which sets the Palmer's method apart from other methods for the extraction of terrain features. The program uses symbolic processing and is compact, but, these characteristics are not particularly meaningful to analysis of terrain features. They are more concerned with the accuracy of results, speed of processing, and ease in using the computer programs. The claimed characteristics of this method do not reflect the user's needs. In fact, Prolog is slow in computation. The extraction of terrain features will require enormous amount of computations and is not suitable for Prolog, which is not designed for calculation. For practical implementation, this method is not appropriate.

The reliability of TINs is a major concern in using this method. On a TIN structure, the valley and ridge lines will be a subset of the segment in the TIN. A comparison of the various configuration of segments shows some are very awkward, it is really doubtful whether the resulting valley and ridge lines will make sense. It has been suggested that information of ridges, valleys, and slope breaks are needed for a faithful creation of TINs. To extract these features from TINs does not appear to be practical so this research will not further pursue the subject. To the contrary, the task of this current research to find valleys and ridges will be beneficial to the construction of TINs.

The strength of Prolog in defining objects and relations makes it very suitable to construct an object-oriented data base of terrain features. For example, a slope unit can be defined by its adjacent ridge and valley lines. A mountain can be built upon slope units. These hierarchical features can be organized in Prolog and carry the inheritance of features. The lines along valley beds and ridge tops are defined at a low level in this hierarchy. The extraction then requires enormous manipulation of the raw data of elevation which is not suitable for Prolog to handle. Yet, when successfully built, it has a great potential to represent a hierarchical structure of terrain features in a way which represents a close approximation of reality.

Chapter 6. Profiling Approach

6.1 Introduction

Geometric properties are essential to human conceptual models of terrain features. The conceptual model of ridges, for instance, would be that a ridge displays a convex pattern along its cross profiles. Conversely, a valley is thought of as having a characteristic concave pattern along its cross profiles. The shape of cross profiles is so natural and intuitive that many researchers have adopted it to develop operational definitions for ridges and valleys. On a regular-grid DEM, ridges are defined as a collection of those points with a convex cross profile and valleys are those points with a concave cross profile.

The shape of a cross profile is a local phenomenon, i.e., it is a property pertaining to a small area. Thus, these algorithms extract terrain features with a local operation, or 'filter' as termed in image processing. The filter will identify a candidate terrain feature within a small area within the DEM, e.g. a 3 by 3 or 5 by 5 window of DEM cells. This filtering process is conducted sequentially to cover the whole data area, as shown in Figure 6.1.

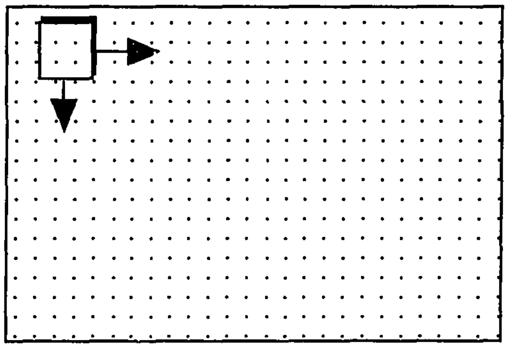


Figure 6.1: A local process.

When the operation runs through the whole area, points extracted from local operations are taken collectively as ridges or valleys. Variations exist in defining 'convex profile' and 'concave profile'. Based on the use of DEM data, the profiling methods can be classified into two groups: discrete approaches and continuous approaches. The discrete approaches take DEMs as the true elevations and analyze them directly. The continuous approaches take the DEMs as elevation data with noise. The latter approaches use DEM data to generate a polynomial to fit the ground surface and analyzes the polynomial instead of the original DEMs. This author identified three profiling methods, including two discrete approaches: discrete profiling and HILO algorithm, and one continuous approach of zero-crossing. In this chapter, the author will dissect their definitions, describe an implementation of each of these methods, and evaluate their performance both individually and generally. Each of the next three sections contains of the analysis of one approach. A discussion section at the end will conclude the performance of profiling methods.

6.2 Discrete Profiling

A direct comparison of elevations along cross profiles is straightforward in meaning and implementation. The discrete profiling approach is built upon this concept. The local operation takes place in a 3 by 3 window of a DEM. Within each window, the elevations along four straight profiles are compared (Figure 6.2). On any of the four profiles, if the central point is higher than its two adjacent points, it will be classified as a ridge point. A central point lower than its two adjacent points on any of these profiles will be classified as a valley point.

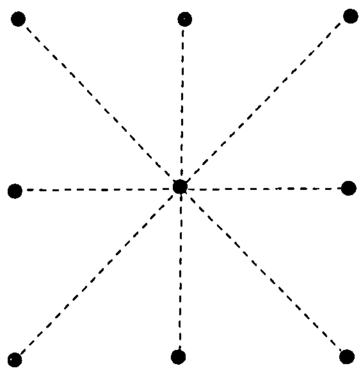


Figure 6.2: Four profiles in a 3x3 DEM window.

This algorithm has been implemented and tested by the author. The Kaneohe data set was used for the test. For each data point inside the matrix, its elevation was compared to that of adjacent points. The extracted ridge points are shown in Figure 6.3 while the extracted valley points are shown at Figure 6.4.

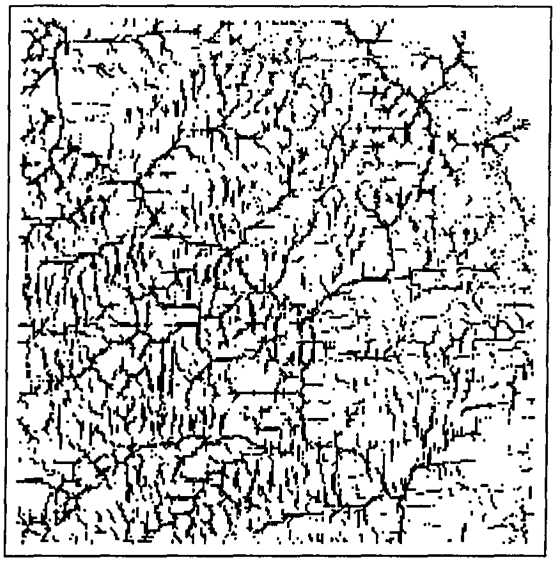


Figure 6.3: Ridges extracted by the discrete profiling approach.

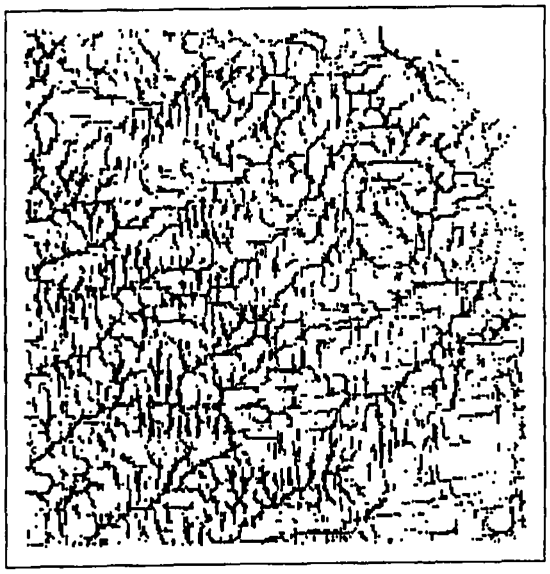


Figure 6.4: Valleys extracted by the discrete profiling approach.

In that test, comparisons were performed without a threshold value relief. The DEM data is composed of integer elevations at a 30 meter spatial resolution. Along the two orthogonal directions, if the central point is only one meter higher than its two adjacent neighbors, the slope will be 1/30, with the center point still being classified as a ridge point. Along the two diagonal directions, a one-meter relief will be 1/51; although much gentler, the center will still be classified as a ridge point according to the definition of this approach. Such definition is rather loose. This loose definition is the cause of extraneous ridge points scattered all over the area. On the other hand, even with this loose definition, there are many gaps on the major structures of ridges or valleys which are supposed to be connected. In a subsequent test, the criterion is raised so that, to be a ridge or valley point, there must be at least a two meter difference between the center and its neighbors. The extracted ridge points are in Figure 6.5.

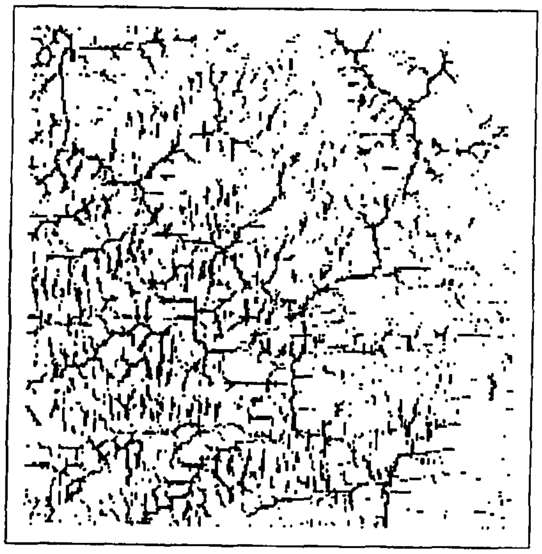


Figure 6.5: Ridges extracted by a 2 m criterion.

This stricter criterion reduces the number of ridge points from 11354 to 7626. It eliminates many scattered points on the low elevation area. However, the main structures of ridges and valleys become even more fragmented. Changing the criterion for the local process will not gain overall improvement. This phenomenon shows a dilemma in keeping the main structures complete while reducing the number of noise points. More analysis in the following sections will reveal whether this dilemma is pertinent to profiling approaches generally.

6.3 HILO (Highest/Lowest) Algorithm

The HILO algorithm (Peucker and Douglas, 1975) takes a 2 by 2 window of a DEM matrix as the working unit. The local operation marks the highest point of the 2 by 2 window. After the operation runs through the whole area moving cell by cell, points that have not been marked are considered valley points. In a similar process of marking the lowest point of the 2 by 2 window, points that remain un-marked are ridge points.

The meaning of this algorithm is not apparent and is susceptible to misunderstanding; further dissection is required in order to comprehend its meaning. For doing that, a 3 by 3 window of DEM with focus on the central point is considered, as shown in Figure 6.6.

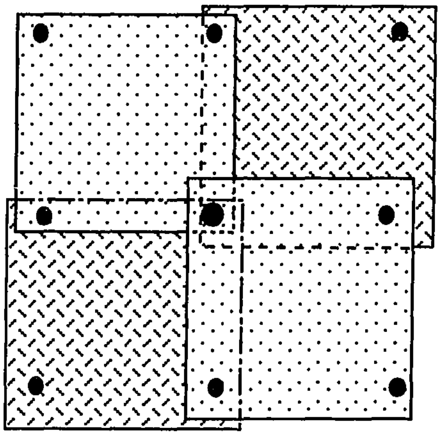


Figure 6.6: A central point within four 2x2 windows.

The central point is involved with four runs of local operations each taking place in the 2 by 2 window at four corners: upper-left, upper-right, lower-left, and lower-right. According to the algorithm, a ridge point must not be the lowest point in each of these four sub-windows, i.e. it must be higher than at least one of the other three points at each 2 by 2 window. This algorithm is similar to the previous discrete profiling approach in terms of the comparison of elevations between neighboring points. However, the discrete profiling method deals with clearly defined profiles while this algorithm does not. Figure 6.7 illustrates several examples of 3x3 window in which the central point is higher than its adjacent points along the cross profile and will be classified as a ridge point.

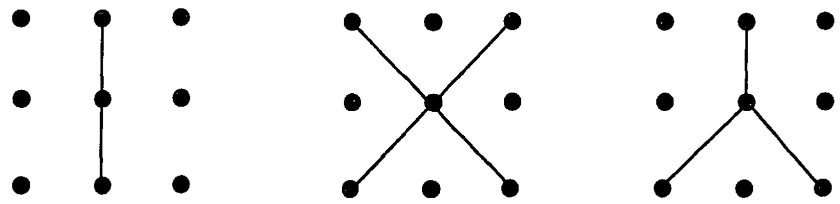


Figure 6.7: Configurations for sample ridge points.

It should be mentioned that there are configurations that are similar to those shown in Figure 6.7 and present a convex shape, yet do not meet the criteria of the HILO algorithm as ridges. The center points of these 3x3 windows may be the lowest point in one of the four 2x2 windows; therefore, they are not extracted as ridge points based on the HILO algorithm. These configurations are illustrated in Figure 6.8.

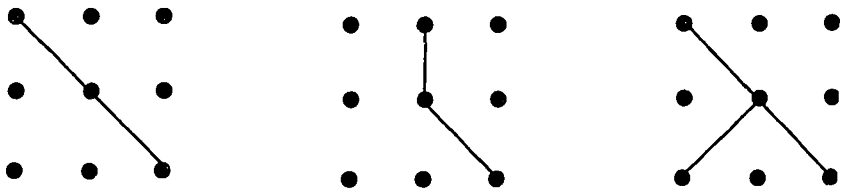


Figure 6.8: Configurations for non-ridge points.

Comparing the conditions in both Figure 6.7 and Figure 6.8, the left most configurations of both figures are similar except one is along an orthogonal profile while the other is along a diagonal profile. These two configurations show how the direction of a cross profile can affect the resulting ridge and valley lines. On DEM data, the directions of ridges and valleys are various and unpredictable. As Figure 6.9 presents, the true ridge on the ground may be along any direction or even between points.

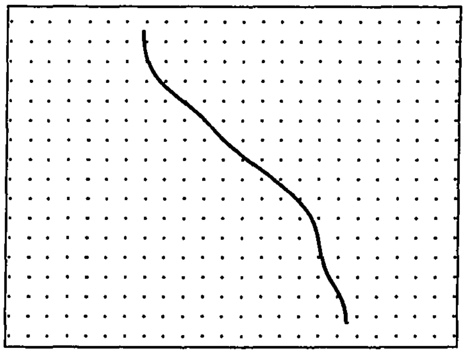


Figure 6.9: A possible ridge line within a DEM.

Using the HILO algorithm, ridge and valley lines along diagonal directions are less likely to be extracted compared to those with a horizontal or vertical direction. This directional discrimination is rather arbitrary and not appropriate.

Compared to the discrete profiling algorithm, the HILO algorithm applies a stricter criterion in defining ‘convex’ and ‘concave’. In a further investigation, the HILO algorithm has been implemented by the author to extract both ridge points and valleys points for the Kaneohe data set. Figure 6.10 presents the extracted ridge points while figure 6.11 shows the extracted valley points.

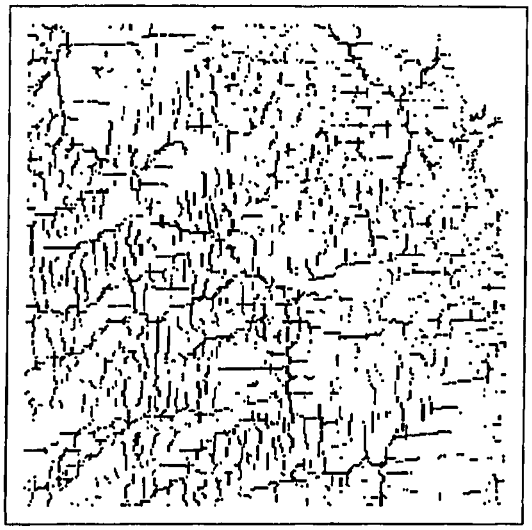


Figure 6.10: Ridges extracted by the HILO algorithm.

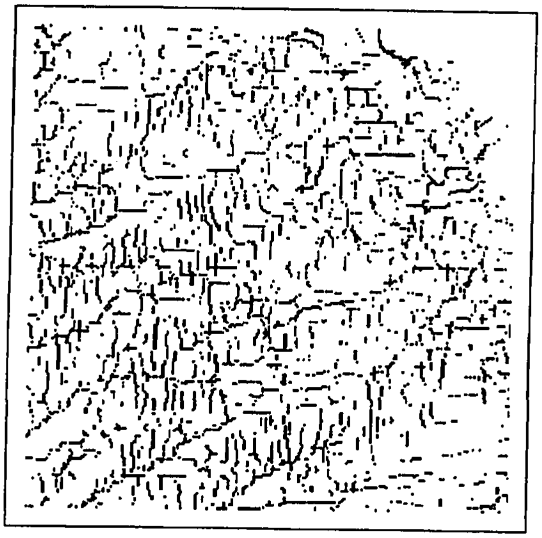


Figure 6.11: Valleys extracted by the HILO algorithm.

This approach extracts many fewer points compared to the aforementioned discrete profiling approach. As the result shows, the main structures are more fragmented than those ridges shown in Figure 6.3. In a further experiment, the criterion of local operation was loosened from 'higher' than at least one point in each 2 by 2 window to 'not lower' than at least one point in each 2 by 2 window. The extracted ridge points are shown in the following figure.

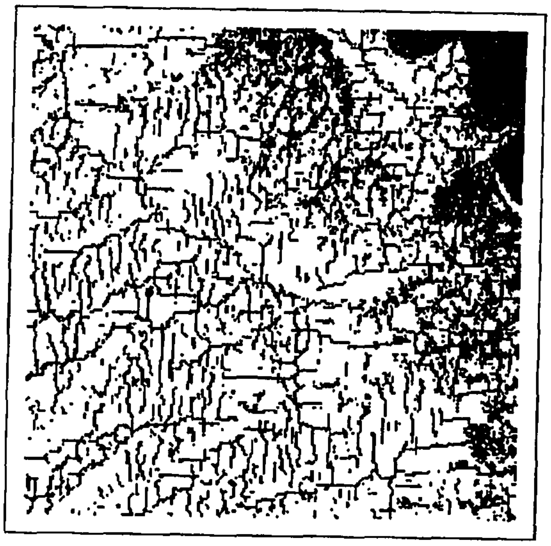


Figure 6.12: Ridges extracted by a loosened constraint.

This criterion extracts many more points and makes the main structure more complete at the cost of including more scattered points and those on a flat area. Again, this test shows that a simple change in the local operation will not result in an overall improvement in extracting features. The dilemma between completeness of main structures and number of noise points pertains to both discrete approaches tested here. In the following section, a continuous approach will be investigated.

6.4 Zero-Crossing of a Facet-Model

Zero-crossing is widely used for edge detection in computer graphics and image processing. A zero-cross occurs when the plot of a mathematical function curve passes the zero line. A zero-crossing may occur to any mathematical function, although the meaning of zero-crossing varies with the function. For example, for a function representing elevation data, the zero-crossing of the first derivative indicates areas where slope is zero. The zero-crossing of the second derivative for the same function indicates that slope is constant.

Haralick (1983) introduced a facet model of zero-crossing to delineate valleys and ridges from remotely sensed imagery. Smith *et al.* (1990) applied this model to identify the same features from DEM data. This method defines a center point as either a ridge or valley point if it is close enough to a zero-crossing along the steepest direction of the embracing local window. Whether the center is a ridge or a valley point can be checked by the rate of slope change. Conceptually, this approach still uses the shape of the cross profile as the only criteria to detect feature points. Nevertheless, this method is technically much more intricate than the two discrete methods described previously. An introduction to the zero-crossing of this facet-method is provided below.

The mathematical equations listed here are to be applied to a 5 by 5 window, the size used by both Haralick and Smith *et al.* In practice, the filtering process and the detection of zero-crossing can be performed in various sizes. For each 5 by 5 window of DEM data, the cubic polynomial in equation (1) is created to fit the original DEM data.

$$f(r,c) = k_1 + k_2r + k_3c + k_4r^2 + k_5rc + k_6c^2 + k_7r^3 + k_8r^2c + k_9rc^2 + k_{10}c^3 \quad (1)$$

$$-2 \leq r \leq 2$$

$$-2 \leq c \leq 2$$

Each coefficient of this equation is unique and needs to be determined from pre-defined fixed masks and the DEM data of the current window. The mask for K_1, for example, is listed below.

-13	2	7	2	-13
2	17	22	17	2
7	22	27	22	7
2	17	22	17	2
-13	2	7	2	-13

For a 5 by 5 window of elevation e_{ij},

$$k_1 = (\ \Sigma \, e_{ij} \times m_{ij} \ ) \, / \, (\ \Sigma \ m_{ij})$$

After creating these coefficients, the steepest slope direction, α, is obtained from the solution of which the second directional derivative of equation (1) equals zero. Thus,

$$\alpha = 0.5\tan^{-1}(k_5/(k_6 - k_4))$$

The value of α with a minimum slope differs from that with a maximum slope in $\pi/2$. From that solution Equation (1) can be converted into a format of polar coordinates by following substitution:

$r = \rho \sin\alpha$ and $c = \rho \cos\alpha$

thus, Equation (1) changes to:

$$f(\rho) = A\rho^3 + B\rho^2 + C\rho + R1 \qquad (2)$$

where $A=(k_7\sin^3\alpha + k_8\sin^2\alpha\cos\alpha + k_9\cos^2\sin\alpha + k_{10}\cos^3\alpha)$

$B=(k_4\sin^2\alpha + k_5\sin\alpha\cos\alpha + k_6\cos^2\alpha)$

$C=(k_2\sin\alpha + k_3\cos\alpha)$

ρ= the distance from the center at the direction of α

Now, the elevation along the steepest direction is a function of ρ. The zero-crossing will occur where the first derivative of equation (2) is zero, i.e.

$$f'(\rho) = 3A\rho^2 + 2B\rho + C = 0 \qquad (3)$$

By solving Equation (3), the zero-crossing will occur at:

$$\rho = \frac{-B \pm \sqrt{B^2 - 3AC}}{3A}$$

If ρ is small enough, i.e., compared to a pre-defined threshold value, the central point will be classified as a valley or ridge, depending on the sign of the second derivative. If the second derivative is positive, the central point is extracted as a valley point, otherwise a ridge point.

A copy of the source code developed by Smith *et al.* (1990) has been received and adapted in this current research. This program has been tested for the Kaneohe data. The extracted ridge and valley points are shown as below.

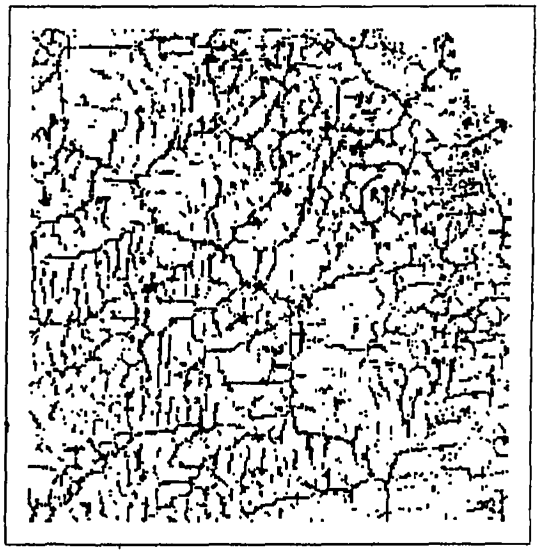


Figure 6.13: Ridges extracted by a zero-crossing method.

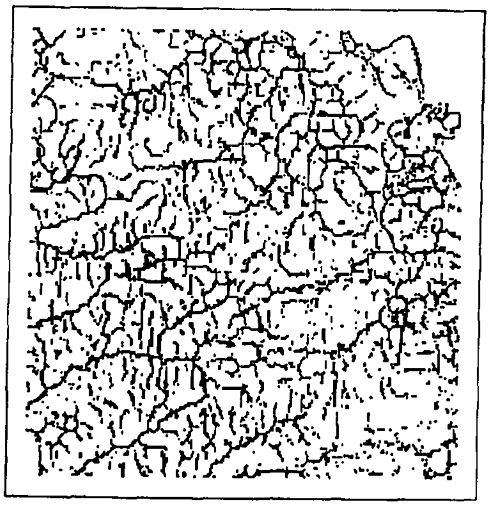


Figure 6.14: Valleys extracted by a zero-crossing method.

By visually comparing these two figures with those figures created from discrete approaches, there is no obviously apparent advantage of this computation-intensive approach. There are still many gaps on the main structures of ridges and valleys. The local noise is still abundant. Besides having the same problem of discontinuity and noise as discrete approaches do, the zero-crossing approach implemented here introduced another problem. The extracted valleys from the zero-crossing method contain many false points, points that do not present a convex or concave pattern in the original DEMs. The former problems of discontinuity and noise are related to the filter size of finding zero-crossing. The later problem is related to the mathematical nature of fitting original DEMs with cubic polynomials. Further investigation of these two aspects of problems follow.

The problems of noise and discontinuity pertain to zero-crossing approaches in general. The usage of zero-crossing techniques normally require a trial-and-error process for determining the filter size and threshold value. If a larger filter is used, the extracted edges will be more continuous, only constituting major structures, thus removing the short features and scattered points. The disadvantage of using a large filter is that it will generalize the major linear features on the data. To show the relations between filter size and the resulting features, the 'i.zc' function, for zero-crossing detection, available in the GRASS 4.0 system is adopted for several tests. Figure 6.15 shows the input image for zero-crossing, a watershed map of the Kaneohe area. Each watershed is homogeneous with a unique number, so the edges between watersheds are clear. The 'i.zc' program was executed several times to find these edges with different filter size and threshold values. Figure 6.16 shows the extracted edges using a filter size of 3 by 3. Figure 6.17 shows results using a filter size 9 by 9. A threshold value of 10 was used by all these figures.

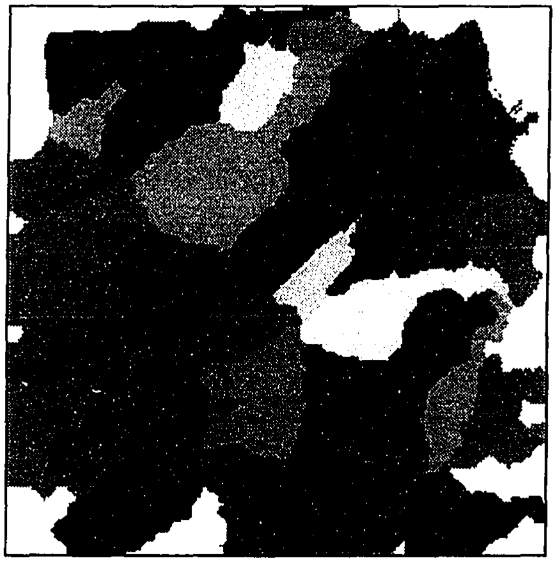


Figure 6.15: An image for tests of zero-crossing.

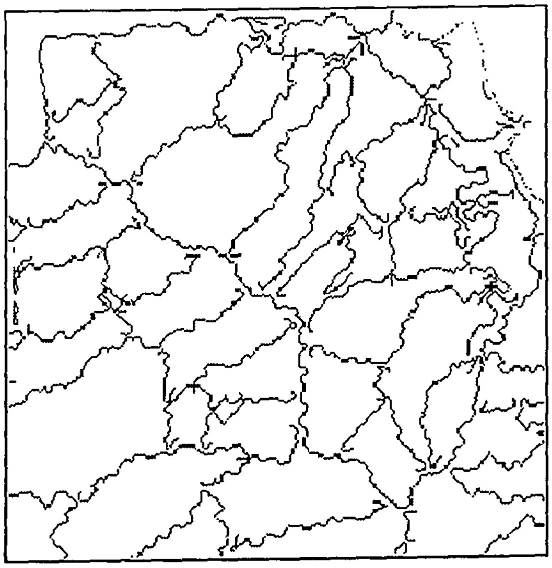


Figure 6.16: Zero-crossing by a 3x3 filter.

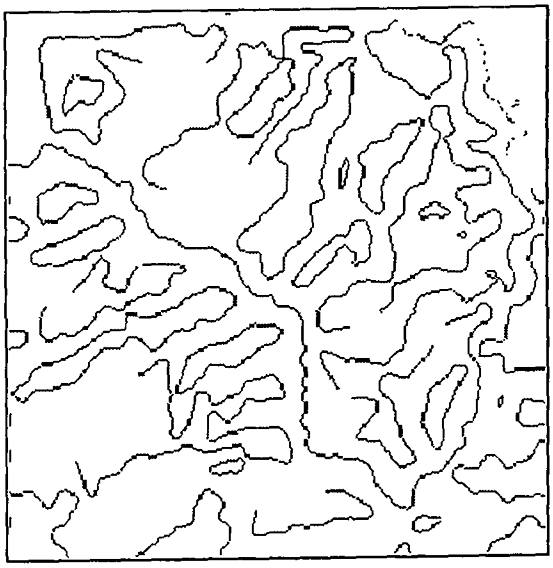


Figure 6.17: Zero-crossing by a 9x9 filter.

As these results show, the filter size greatly affects the detail of the results. With a 3 by 3 filter, the resulting edges are most complete. The 9 by 9 filter extracted many fewer edges, as Figure 6.17 shows, and the resulting edges were smoother than those extracted from a 3 by 3 filter. In this test, the 3 by 3 filter is better than filters of other sizes, due to the fact that the watershed image is composed of homogenous polygons thus the edges are clear. For an image with graded variation on its pixel values, the selection of filter size will be very difficult. Partly to investigate the filter size problem further, but also to test the feasibility of the concept that ridges are boundaries of various slope units, more experiments with zero-crossing were conducted using the 'i.zc' command of GRASS4.0.

The slope aspect of Kaneohe is used as source image for finding zero-crossing, as shown in Figure 6.18.

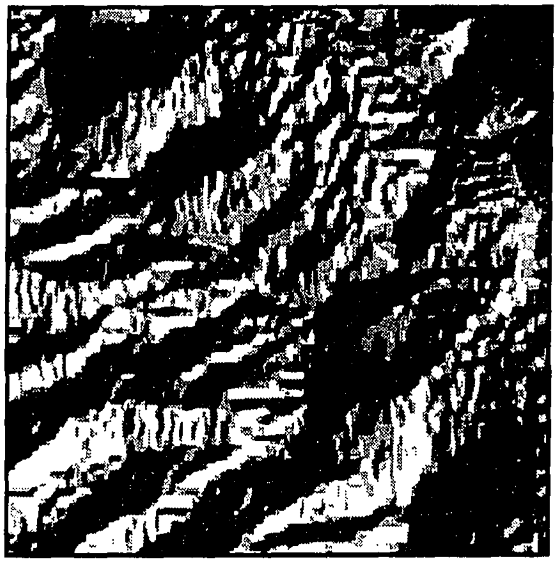


Figure 6.18: An aspect map of Kaneohe DEM.

The values in the aspect map, ranging from 0 to 25, indicate the slope aspect of this area. The value 0 means no data, value 25 means flat areas, each value from 1 to 24 represents a specific aspect category at a 15-degree interval. As shown in Figure 6.18, the boundaries of slope suggest the locations of many ridges, yet it is questionable whether we can delineate the designated boundaries as ridge lines. In a sequence of experiments, filter sizes of 3, 9 , and 15 were used respectively for finding zero-crossings. The extracted edges are shown below.

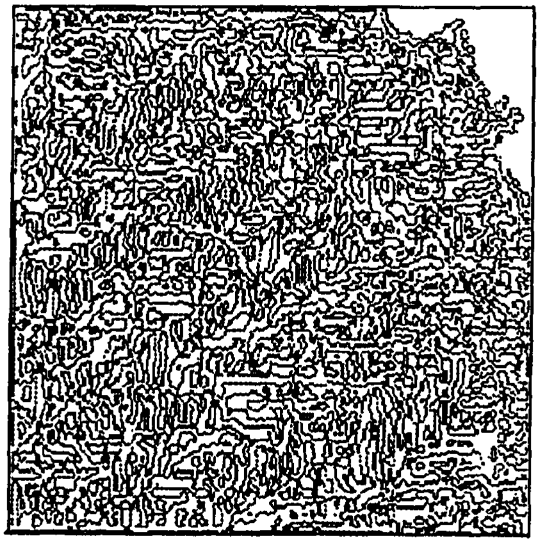


Figure 6.19: Zero-crossing by a 3x3 filter.

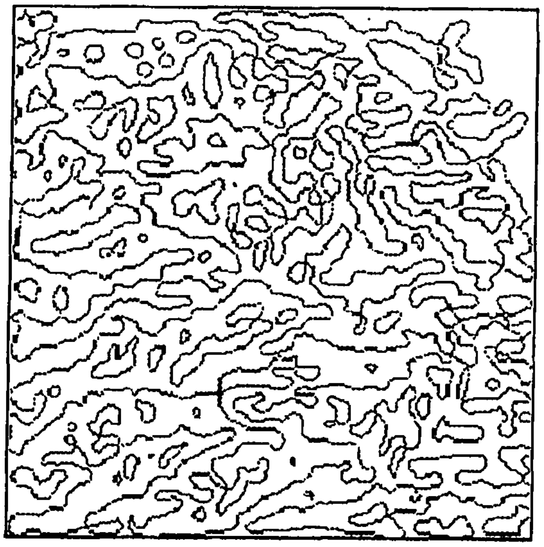


Figure 6.20: Zero-crossing by a 9x9 filter.

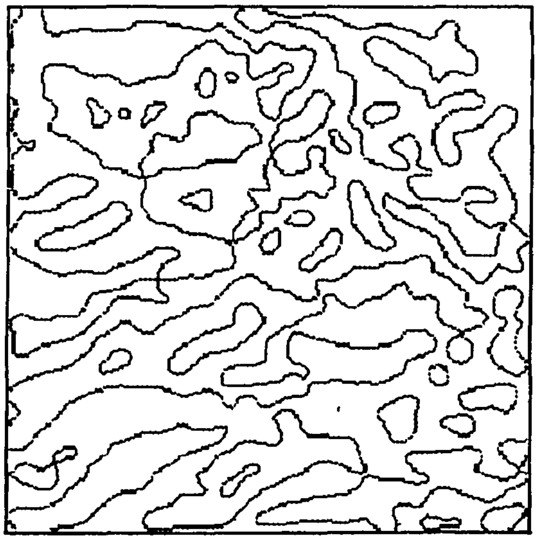
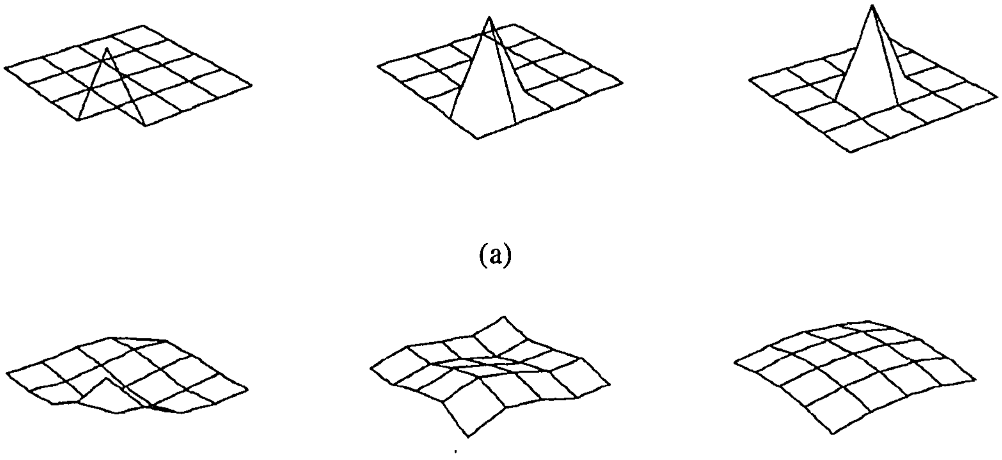


Figure 6.21: Zero-crossing by a 15x15 filter.

These figures show that various sizes of filters extract edges at different levels of detail. The 3 by 3 filter extracts many small features and misses the major structures. The 15 by 15 filter delineates the boundaries between major blocks of slope units. If edges of various level are designated, it is necessary to run the computer program several times with different filter size to find zero-crossings and integrate them together. These processes involve much trial-and-error, which are tedious and difficult to automate. For the development of automatic extraction of terrain features, these complicated zero-crossing techniques do not provide any clear advantages to compensate for their being tedious and expensive in computation.

The fitting process of creating a cubic polynomial for each 5 by 5 window raises some concern. An examination of Figure 6.14, showing the extracted valley points by the zero-crossing approach, finds a series of 'valley points' along the coastal line at the upper right corner. These points suggest that there is a continuous concave pattern along that area. However, on the original DEMs, there is no such pattern. Apparently, the fitting process of the facet model introduces inflections which produce fictitious valley points. Such false inflections are common in a cubic polynomial. For a further investigation on such inflections, several tests were conducted using S-plus, a computer package for interactive data analysis and display. This author input several 5x5 matrices into S-plus, implementing the filtering function inside S-plus. The original matrices and the resulting surface from filtering are shown in Figure 6.22.



(a)

(b)

Figure 6.22: An illustration of fitting problem of a facet model, (a) original surfaces; (b) resulting surface by the facet model.

These tests show some possible inflections introduced by the filtering process. In a continuous surface, these inflections may be part of the surface. For the representation of ground surface, however, these artificial inflections are rather unnatural and misleading for the subsequent analysis. If the filtering results fail to reflect the true characteristics of the original surface, the result of zero-crossing will be questionable.

In conclusion, the zero-crossing method is complicated and problematic. Two problems have been identified - filter size and fallacious inflection. There is no apparent resolution to these concerns. Furthermore, the results from zero-crossing is not any better than results from discrete approaches. The approaches such as zero-crossing continuous may be theoretically intriguing, but not practical for the extraction of ridge and valley lines.

6.5 Conclusion

This chapter analyzed and tested three profiling algorithms for the extraction of terrain features. At a micro level, there are differences between feature points extracted by these three approaches. Since each of these approaches adopt different criteria in defining features, points extracted by one approach do not necessarily appear in the results from other approaches. In general, the results from all approaches were very similar: disconnected and noisy. This similarity resulted from the adoption of a local process to terrain feature extraction.

Conceptually, the geometric pattern of cross profiles is sufficient to define ridges and valleys. Throughout the experiments in this chapter, the author concluded that geometric patterning itself is both difficult to implement and insufficient to define terrain features. Geometric properties are very sensitive to the data resolution. Interpretation of shape is closely affected by the scale of area examined and the resolution of the terrain model. To be able to extract terrain features of various magnitude, the criteria used for defining shape should be flexible enough to deal with various sizes. The algorithms tested in this chapter all adopt a fixed scale to interpret the terrain data and consequently miss features at other scales. In the discussion of the tracing process, it also becomes apparent that shape alone does not completely define terrain features. Elevation, length, and spatial location of an entity with continuous pattern of 'convex' or 'concave' affect whether or not such entity will be classified as a given terrain feature.

The profiling approaches provide a general location and pattern of the ridges and valleys. Yet, before further development to fix the continuity and noise problem, they are not suitable for practical purposes. Technically, finding the convex pattern is not as easy as it appears to be. Conceptually, the geometric pattern of cross-profile is not critical enough to separate features and non-features. To make these profiling approaches useful, further research is needed. Smith *et al.* (1990) suggested a tracing process to construct a connected drainage network. Several other tracing approaches were also developed by different research groups. Next chapter will investigate the performance of these tracing approaches.

Chapter 7. Tracing Approach

7.1 Introduction

The preceding chapter investigated several profiling approaches. It is concluded that the resulting features from those profiling approaches are discontinuous and contain too much noise, and thus are hardly useful for practical applications. The discontinuity and noise in these resulting features are inherent from the adoption of local and segmented operations on DEMs. These undesirable characteristics are difficult to fix by improving the local operations; therefore, alternative approaches are sought. Riazanoff *et al.* (1988) introduced three algorithms using a feature tracing procedure for the extraction of terrain features. Smith *et al.* (1990) also adopted a tracing process to fix the segmented features resulting from a zero-crossing approach. In distinction to the segmentation of profiling approaches, tracing approaches will delineate terrain features continuously. As a result, the extracted ridges/valleys are supposed to be better connected with fewer short and scattered features.

In fact, tracing is not new. Manual delineation of valleys and ridges is a tracing process. To extract a ridge line, for instance, human interpreters may first identify a peak as a starting point. From a peak, they delineate the ridge line by tracing along the direction displaying a convex cross profile until there is no more convex pattern along the profiles. The manual extraction of valleys follows a similar procedure. Although the automatic tracing approaches are carried out in computers and deal with DEMs, their problem-solving procedure is similar to that of the manual process. These automatic tracing methods consist of two steps: 1) the identification of starting points; and 2) the tracing along DEMs based on a progress constraint that defines the direction of progressing. The starting points are normally selected from one pass of filtering of the whole DEM matrix and are stored in a list. With a computer, rather than drawing a line, the tracing process can be thought of as a search among the adjacent neighbors conducted at the end of a current feature line. The constraint of each tracing approach will determine which neighbor(s) should be added to the current features. The newly selected point(s) will in turn become an end point for further tracing. The tracing process will continue until 1) the traced line reaches the border of the DEM matrix, or 2) the traced line meets another existing feature line; or 3) no adjacent points of the current end point meet the criteria of progress constraint. There are various ways of selecting starting points and setting progress constraints, which distinguish tracing approaches from each other.

This chapter investigates the performance of automatic tracing approaches. Besides providing potential alternatives for terrain feature extraction, this investigation is intriguing for two reasons. Firstly, the automatic tracing approaches are close renderings of the manual extraction of terrain features. This investigation provides comparisons between extraction by computers and by humans, as well as between characteristics of DEMs and contour maps. Such comparisons will lead to a better understanding of the characteristics of DEMs and computer methods, and better insights for further developments. Secondly, after reviewing the previous work, it is anticipated that the resulting features from these automatic tracing approaches will be very different. Riazanoff *et al.* (1988) linked the different approaches with various applications, e.g., the streaming approach is suitable for hydrological research. Yet, to this author the adequacy of such linkage deserves further scrutiny. The difference of results may reveal a variation in the definition of terrain features; on the other hand, it may suggest a fallacy embedded in the implementation or the nature of tracing approaches. Only after identifying the reasons and sources of difference, can one determine whether such linkage is suitable and desirable. The next section introduces the algorithms of tracing approaches. After that, a detailed analysis and investigation will follow.

7.2 Procedures of Tracing Approach

The three tracing algorithms introduced by Riazanoff *et al.* (1988) are the streaming algorithm, the walker algorithm, and the main saddle points algorithm. In common, these algorithms treat valleys and ridges as the duals of each other. By inverting the criteria of starting points or progress constraints, e.g., descending rather than ascending, the same algorithms for the extraction of ridges are applicable for the extraction of valleys. These three algorithms are described below.

- The streaming algorithm

The tracing of the streaming approach starts from a saddle point. Conceptually, a saddle point is a low point on a ridge, or the intersection of ridges and valleys. On a contour map, such saddle points are normally surrounded by two pairs of contour lines at two orthogonal directions, as shown in Figure 7.1.

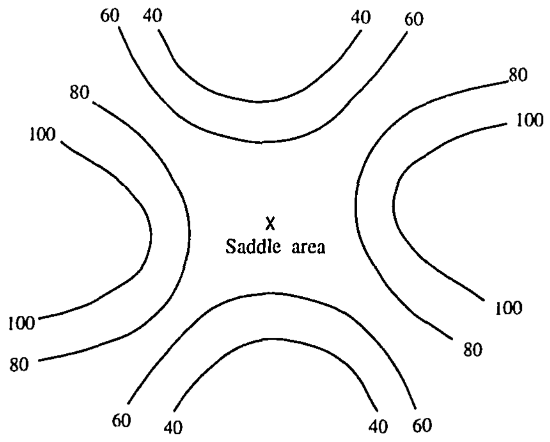


Figure 7.1: An illustration of a saddle point.

Applying this concept into DEMs, Riazanoff *et al.* defined the saddle point as a point with at least two groups of contiguous neighboring points lower than it and two groups of contiguous neighbors higher than it. Figure 7.2 shows several samples of 3x3 DEM window which are considered as saddle points, based on this definition.

+ - +	+ - -	- + -	+: higher than center
+ P +	+ P +	+ P +	-: lower than center
- + -	- - +	- + -	P: center point

Figure 7.2: Samples of saddle points defined by Riazanoff *et al.*

Starting from a saddle point, to extract a ridge line, the progress constraint of this algorithm is to 'climb along the steepest direction'. The same set of saddle points is applicable for the extraction of valley lines with a progress constraint to 'descend along the steepest direction'.

- The walker algorithm

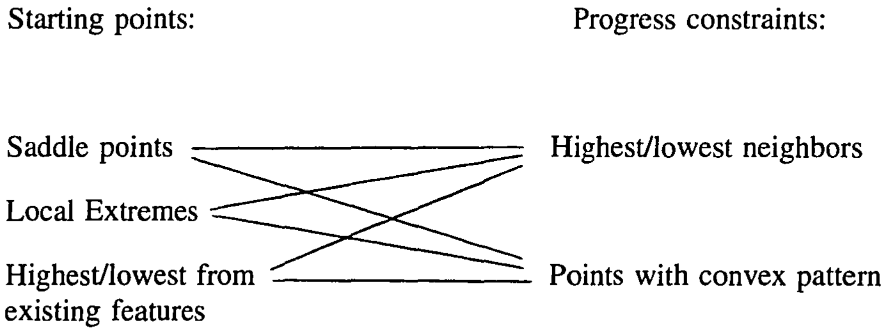
For extracting ridges, this algorithm starts from a local maximum: a point higher than all its neighbors. The progress constraint is to 'trace along those points posing a convex pattern along their four profiles less the one from the existing feature'. The extraction of valleys starts from local minima, or pits, and traces along those points carrying a concave cross profile.

- The main saddle algorithm

The selected points of this algorithm are 'main saddle points', as named by Riazanoff *et al.* A main saddle point is a saddle point that is not on the network extracted from the aforementioned Walker algorithm. The progress constraint of this approach is, identical with that of the Streaming algorithm, to climb along the steepest direction for ridges and to descend along the steepest direction for valleys.

Besides Riazanoff *et al.*, Smith *et al.* (1990) also derived a tracing approach as a follow-up operation of their zero-crossing approach. Working on the resulting features from a zero-crossing approach, their tracing process first selects all those features that are longer than a pre-defined threshold length. Subsequently, the highest point of each of these selected features is identified and stored into a list. The list is then sorted according either to the elevation of each point or to the length of the feature from which each point was selected. These points are the starting ones for the tracing process. The progress constraint is to trace along the lowest neighbor, with distance adjusted, until the current feature reaches a pit, border, or already extracted feature.

Although only four tracing approaches were described above, more methods can be derived from these four. As a generalization, those four approaches employed three types of starting points with two types of progress constraints. By different combinations of those alternatives of starting points and progress constraints, six different tracing approaches should be available, as shown below.



By adding a new type of starting point or progress constraint, the number of automatic tracing algorithms may even be greater. Instead of trying to test these tracing algorithms one by one, this research takes a more fundamental approach by investigating two components of the tracing algorithms: selection of starting points and progress constraint. The performance of each tracing method relies on the feasibility and adequacy of the two steps of tracing -point selection and tracing.

7.3 Analysis and Experiments

There are theoretical and technical problems related to these two tasks. From a theoretical aspect, this author is concerned with whether the starting points and progress constraints are sufficient and necessary for the delineation of valley and ridge lines. From the technical aspect, the concern is whether the implementation of these algorithms can extract those designated points and terrain features. The following section will first investigate the selection of starting points and progress constraints separately. Thereafter, a general evaluation of tracing approaches will be provided.

7.3.1 Selection of Starting Points

The selection of starting points critically affects how many and where the extracting features will be. Each selected starting point indicates the existence and beginning of a feature line. The starting points designated by Riazanoff *et al.* are saddle points, peaks, and pits. These points are terrain features important to the study and representation of terrain. Similar to the extraction of valley and ridge lines, the extraction of these points may be problematic, involving the complexity in their meanings and scales. This investigation takes a practical and experimental approach to investigate the extraction of these points. A sequence of tests has been conducted to extract the saddle points, local maximum, and local minimum for the Kaneohe DEM.

The saddle points were extracted using the definition by Riazanoff *et al.* The extracted points are given in Figure 7.3.

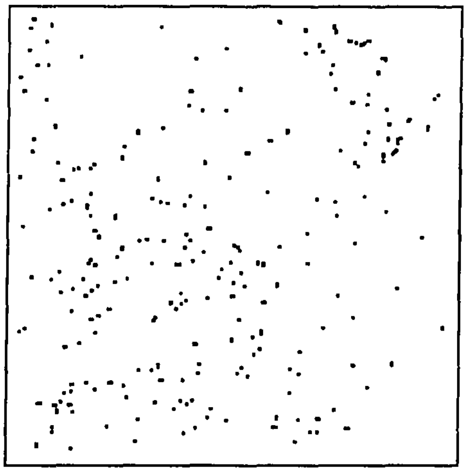


Figure 7.3: Saddle points extracted by the definition of Riazanoff *et al.*

From a different perspective, saddle points are thought of as the intersections of ridges and heads of valleys. Thus, in a further test, saddle points were defined as those that were extracted as both a ridge point and a valley point by the profiling method. The result is shown in Figure 7.4.

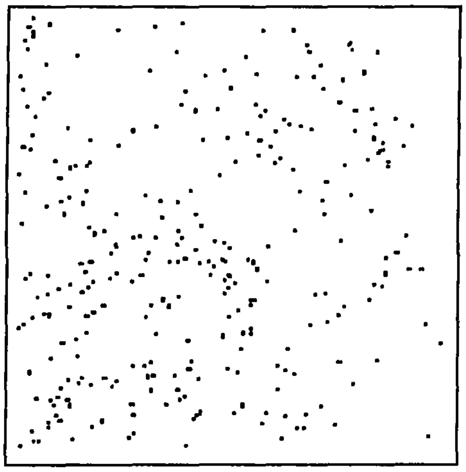


Figure 7.4: Saddle points extracted as intersections of ridges and valleys.

The preceding method extracted many fewer points compared to those extracted by the definition of Riazanoff *et al* (1988). Apparently, by comparing these two different results, there is no single answer to the selection of saddle points. Another concern of this result is that some saddle points are located in areas where ridge valley lines are normally not expected, i.e., flat plains.

The meaning of local peaks and pits are more straightforward, since a peak is higher than all its neighbors. By the same token, a pit is lower than all its neighbors. These characteristics were used as the working definitions for the following extraction of peaks and pits. The resulting points are shown in Figure 7.5 and Figure 7.6 respectively.

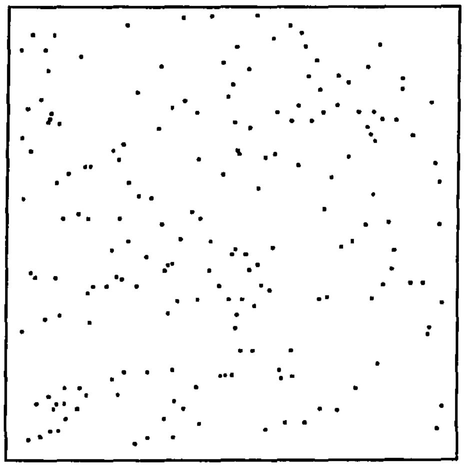


Figure 7.5: Local maxima.

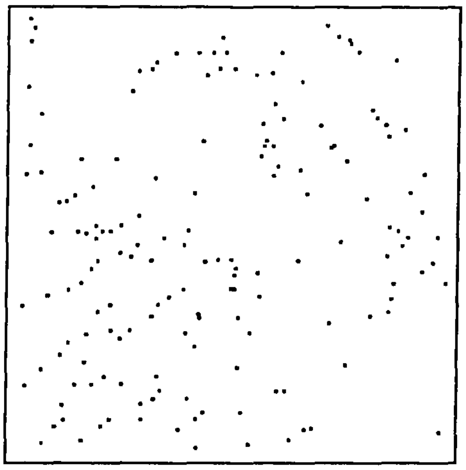


Figure 7.6: Local minima.

By referring these points to elevation, it is apparent that not all these points of local extremes are desirable for the extraction of valley and ridge lines. For example, many 'peaks' are in the low elevation and flat areas where one normally does not delineate ridges. On the other hand, by observing the elevation image, peaks exist in areas where the extraction did not extract even one. These problems of missing points and inclusion of undesired points are caused by the same reasons pertaining to the discontinuity and noise of the profiling methods. Similar to the local processes of profiling methods, the extraction of these points was confined to a small area, a 3 by 3 window of DEM matrix here. Thus, these extraction approaches fail to distinguish significant peaks from those unintended points of minor relief. The discrete nature and integer elevation of DEMs also cause problems. A peak on the ground may not meet the criterion of being higher than all its neighbors on a DEM, due to the truncation of integer elevation. In other words, the true peak is situated between points of a DEM and thus fails to show its height, as following figure shows.

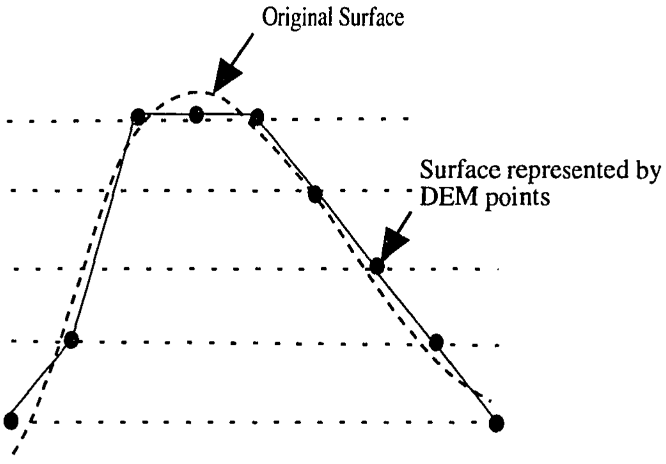


Figure 7.7: An illustration of missing peaks on a DEM.

The starting points selected by Smith *et al.* are not apparent nor clearly-defined on the ground surface. Rather, they vary with the threshold value set by the user. This process of point selection has been implemented and tested on the Kaneohe DEM. The resulting points are shown in Figure 7.8.

Figure 7.8: Starting points selected by Smith *et al.*'s approach.

By viewing the distribution of these resulting points, this author noticed some problems inherent in this process. Firstly, the length of a feature alone may not be sufficient for the selection of starting points. For example, a feature line located at a high elevation is very likely to indicate a start of a valley line, compared to those located on low and flat areas. Secondly, this extraction process will result in fewer points for well-connected features while extracting many more points from the fragmented features, given that each feature is longer than the threshold value. The consequence is that some valley lines may miss their starting points and thus be excluded from the subsequent tracing stage while many feature lines may be extracted from the fragmented features.

Previous tests and analyses reveal the difficulty and ambiguity in selecting the starting points. These problems undermine the feasibility of the tracing methods in general.

7.3.2 Progress Constraints

Two types of progress constraints have been identified: tracing along the steepest direction and tracing along points with convex/concave cross profiles. In an idealized condition, either of these constraints should be sufficient to delineate a ridge and valley lines. For example, starting from a saddle point, a typical ridge should be along the highest elevation, which should also pose a convex pattern along its cross profile, with each small step of movement. However, it is questionable whether these two constraints are generally applicable to the extraction of ridges/valleys of various types and scales.

The tracing approach, based on the criterion of cross profile, will only extract those points with a concave/convex cross profile. The resulting features from this tracing approach is therefore a subset of those extracted by the discrete profiling method that was discussed in the previous chapter. This tracing approach is in fact the discrete profiling approach with an extra clearance stage. The tracing process starts from a selected point, thus line features containing no starting points will not be traced. This criterion will remove many short and scattered features from the result. However, the major feature lines will remain disconnected since this tracing method terminates when there is no neighbor of current end point displaying a convex/concave profile. Thus, the progress constraint based on cross profile cannot extract a connected and complete network of ridge and valley.

Tracing along the steepest direction, or lowest/highest neighbors adjusted with distance, is a rather loose constraint. Most points on a DEM have lowest and highest neighbors. Therefore, this progress constraint will result in a connected network. A problem with this constraint is that the highest/lowest neighbor may not necessarily be qualified as a ridge or a valley point. This problem can be better illustrated by a sample contour map and its corresponding sample DEM, as shown in Figure 7.9.

142	142	142	142	142	142	142	142	142
130	130	130	130	130	130	130	130	130
119	119	119	119	119	119	119	119	119
105	105	106	107	109	108	106	105	105
89	90	92	94	96	98	96	94	90
72	80	84	88	90	92	89	86	80
65	70	75	82	84	83	80	79	75
60	63	68	74	78	77	76	71	65
50	56	62	68	79	69	66	62	55
45	48	54	62	63	61	57	55	48
40	42	48	56	58	53	51	47	42
40	33	33	57	56	53	50	50	33
35	40	43	51	50	48	42	40	38

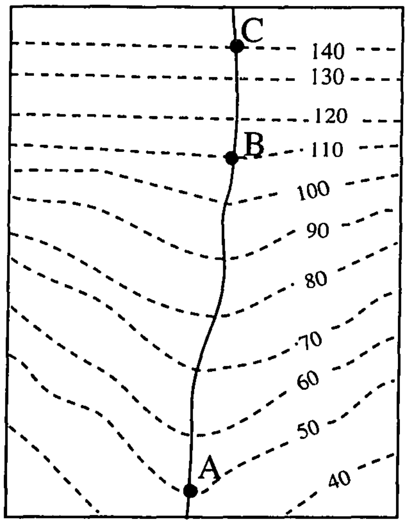


Figure 7.9: Highest points are not necessary along ridges.

Moving from point A to point B, on both the contour map and the DEM, the tracing along the steepest slope direction delineates a ridge line. From B to C, the tracing process has no problem in climbing along the steepest direction. Yet, the delineated line is not normally thought of as a ridge line; it does not pose any convex/concave pattern along its profile. Here, tracing along the highest direction does not result in a ridge line, but only the slope line passing point B. Unfortunately, this type of situation is not rare. Fallacy can occur when a ridge merges into a homogeneous slope unit or the tracing process derails from the true paths of ridges. The progress constraint of steepest direction overlooks these conditions and likely to extract false terrain features.

7.3.3 Integration of Whole Process

After identifying the potential flaws and problems in each of the two steps of tracing, this author is curious about the performance of these tracing procedures. As mentioned in previous section, the progress constraint of cross profile will result in disconnected features no matter what type of starting points one uses. Thus, no experiment will be conducted to test those tracing approaches adopting cross profile as progress constraint. Instead, the progress constraint of tracing along the lowest/highest neighbor will be tested with various types of starting points.

In the first test, the local maxima of Kaneohe DEM were selected as starting points for tracing. For each starting point, the computer will trace along highest neighbors at the two opposite directions. This approach reflects the concept that if one is standing on top of a peak and moves along the least relief, one will remain on a ridge. The resulting ridge lines are shown in Figure 7.10.

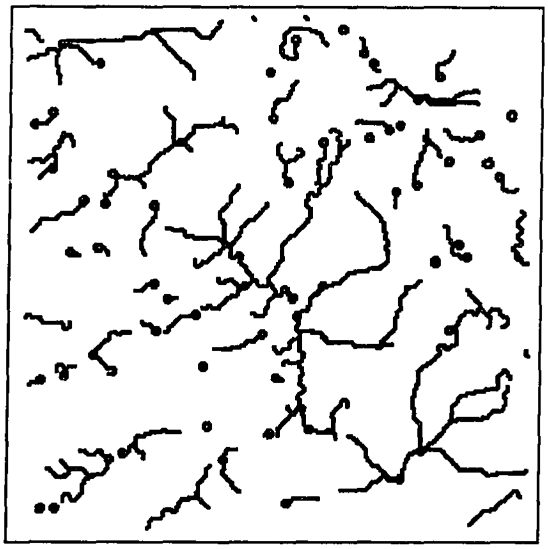


Figure 7.10: Ridges extracted by tracing started from peaks.

In the second test, the saddle points of Kaneohe DEM were selected as starting points. By tracing along the highest neighbors, this test aimed at delineating ridges. The resulting ridges are shown in Figure 7.11.

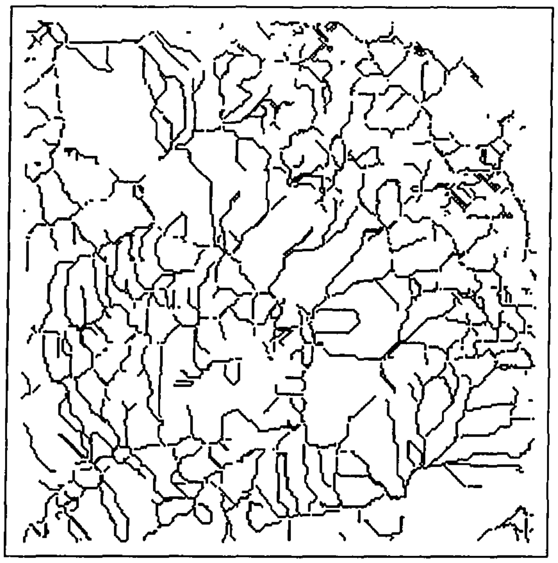


Figure 7.11: Ridges extracted by tracing started from saddle points.

By observing these two results, several problems have been found. First, these two results are very different. Second, the resulting features present loops which are not common in the real world terrain and do not pertain to the ground surface of this test area. Several resulting ridge lines are straight. Through cross-referencing with the DEM data, many of these lines are slope lines rather than ridges. These various problems confirm the concern of early discussion on the points selection and tracing processes.

7.4 Discussion

The results from tracing approaches are very different from those from profiling and zero-crossing approaches. In general, these tracing approaches result in a better connected pattern of features with many fewer noise and short features. Compared to the results by profiling methods, the tracing approaches show a progress toward a better rendering of manual extraction. On the other hand, as shown in the resulting features obtained by this author, and those results by Riazanoff *et al.*, different tracing approaches extract terrain features very differently, which is reflected in the amounts and locations of the extracted features. Such difference has been addressed by Riazanoff *et al.* who linked the difference with various applications. Such linkage, without first identifying the source of difference, is questionable.

Each of these tracing approaches requires a list of starting points. The adoption of certain types of starting points, for which many alternatives exist, is a major source of difference. However, not every ridge and valley line contains these various types of starting points. Peaks, for example, are one option of starting points for the extraction of a ridge, although from a conceptual aspect as well as a technical aspect, not every ridge contains a peak point. Conceptually, the top of a ridge may be somewhat flat so one does not identify a peak. This is particularly true when one is only dealing with a small part of a relatively large network of ridges. Technically, as mentioned above, the peaks on the ground surface are not necessarily presented in the DEMs, because of the truncation and discrete character of DEMs. In either case, the delineation of ridges starting from peaks will result in the loss of ridge lines. Similarly, there are ridge and valley lines that do not contain saddle points either conceptually or technically. The tracing approaches that start from saddle points will miss some other ridges too. Using the saddle points or local extreme as starting points will introduce bias in identifying certain types of ridge and valley lines. A ridge line without a local peak will not be extracted by the streaming algorithm. Similarly, a ridge line without a saddle point will not be extracted by the walker algorithm. Such bias is arbitrary and not justified. It reflects the erroneous aspect and incompleteness of these approaches. Furthermore, on a discrete data set, if one descends along the direction of least slope, stops at one point, and then climbs back along the highest neighbors, the two paths one obtains may not be identical. In other words, ‘ascending’ and ‘descending’ are not directly reversible in the tracing process. Consequently, tracing in the DEMs with different starting points is very likely to result in different terrain features. This author thinks such difference is more a weakness than a desirable characteristic.

The manual tracing approach usually results in connected and complete networks, yet the same procedure implemented on computers with DEMs may not produce a satisfactory result. The difference between manual and automatic approaches can be attributed to two factors. First, the representation of terrain is continuous on a contour map while discrete on a DEM. When searching for the steepest direction, for instance, it is always possible to follow the true steepest direction on a contour map. However, the same tracing process on DEMs is limited to a fixed number of directions, mostly eight, and is very likely to lose the track of true ridge. Second, a human interpreter can flexibly adjust the size of working area on contour maps during the delineation process while most automatic methods are not as flexible. A ridge network is normally composed of ridges of various magnitude and slope. Human interpreters can adjust their criteria to deal with ambiguity or uncertainty arising from these various scales and shapes. In comparison, most of the developed automatic tracing procedures work on a fixed size of DEM matrix with rigid criteria. Flexibility is needed to make the automatic tracing approaches more successful.

The performance of profiling approaches and tracing approaches are complementary with each other in two respects. First, the profiling approach generally identifies most of the terrain features with many short noisy segments. The tracing approaches, on the other hand, produce fewer but continuous points of terrain features. Second, the performance of profiling methods is more consistent and reveals the general location of the terrain features, while the tracing approaches are less reliable in terms of the accuracy of the resulting features. A local error of DEM data or a minor relief on the ground may cause serious misleading effects in the tracing results. It is common for DEMs to contain minor relief in a local area, either due to error of DEMs data or true minor relief on the ground. For a profiling method, local relief will affect only the extraction of resulting features around a small area without causing major change. For the tracing process, however, the local relief may derail the search and consequently change the delineation. Therefore, the tracing process is very sensitive to minor relief of the local topography.

Both profiling and tracing approaches have their advantages and drawbacks, while their performance is complementary to each other. For future development, an integration of these two groups of methods should be promising. The integrated approach should make use of the advantages of each group of methods. It may start with a profiling approach to identify all the potential candidates of ridge or valley lines and be followed by a tracing process to further process the features resulted from the profiling process. In terms of adoption of starting points, the tracing process should trace all segments with several factors of concern. The progress constraints should include the lengths of segments, their elevation and slope, and their adjacency to major features lines. The proposed integration of profiling approach and tracing approach is only an outline. Substantial progress in spatial reasoning is needed before a possible implementation of the proposed method.

This current research has investigated three groups of methods for the extraction of ridge and valley lines. There is one more group of methods based on a hydrological perspective. The next chapter will investigate the potential of this group of methods.

Chapter 8. Hydrological Approach

8.1 Introduction

The development of the hydrological approach, initially concerned with the extraction of stream channels, was based on the fact that channels are locations where overland flows accumulate. This approach adopts the regular-grid structure of DEMs as the working terrain model. By tracing the direction of overland flow and computing the accumulation of overland flow of each cell, the stream channels of an area can be extracted. Although this method does not explicitly adopt the property of shape in the definition, implicitly, shape is still a property pertaining to channels extracted from this method. The flow direction of each cell is determined by its slope and aspect which are characteristics of shape. With a screening based on drainage accumulation, eventually those cells with accumulation exceeding a threshold will be locations where overland flows of two opposite slopes emerge. So this definition of stream channel reflects the conceptual model that stream channels are the intersection of two opposite slopes and that stream channels often carry a V shape profile.

The hydrological approach requires intensive computation and would not be feasible without the computing power of modern computers. In fact, Smith *et al.* (1990) commented that this method was slow because of the enormous computation required. However, physical computational limitations are decreasing. The huge computation no longer presents a constraint to this approach.

The hydrological method has been implemented in the GRASS software version 3.1 and version 4.0. Each version provides a function to delineate drainage networks and watersheds. These extraction functions were developed by different authors and their performance and user interface are different as well. The following section introduces the built-in procedures of these functions of channels extraction.

8.2 Procedure of Stream Extraction in GRASS

The stream extraction function in version 3.1 is written by Lammers and Band (1990) and named 'watershed'. This program involves six steps. Each step requires users to respond to various prompts. The functions and purposes of these six steps are explained below.

1). Filtering the elevation data

This process smoothes the elevation data with a special filter. The filtering is done in a weighted average of 3 by 3 window of DEM. The purpose of this process is to reduce the number of pits which will be identified in the next step.

2). Locating the pits

Pits are points with an elevation lower than all their neighbors. Those overland flows which run into a pit will not be able to find an outlet and cause a termination of stream flow. Although some pits may be part of the terrain relief which form a lake, most of the pits found in a DEM will be due to data resolution and errors and should be fixed. In order to avoid the termination of a stream channel, it is necessary to identify pits and treat them with special care. When water flows into these pits, the elevation of the lowest neighbor is used instead, so that the water flow can find an outlet and continue.

3). Calculating drainage accumulation/outlining watershed

This step is to calculate the drainage accumulation values from elevation data. Drainage accumulation is designated by assigning each cell a value equivalent to the number of cells which drain to it. The example in Figure 8.1 shows the relationship between elevation and accumulation value.

257	246	233	228	240	258	276	298	307	309	306	307	310	309	304	299
259	241	225	223	236	253	269	288	295	293	287	287	290	291	289	287
260	236	216	216	232	249	263	279	283	278	269	267	269	272	275	274
252	228	212	215	229	244	256	266	268	262	253	251	253	257	262	262
236	219	210	216	227	239	246	251	251	246	239	238	241	246	251	252
219	209	209	217	226	234	237	237	233	230	226	225	229	235	240	242
202	200	208	219	225	228	228	222	216	214	212	212	217	223	229	231
197	197	206	215	220	223	222	217	211	210	211	211	215	222	227	229
195	197	204	211	215	217	218	214	209	209	213	214	214	222	226	227
193	196	201	206	209	212	213	210	207	208	215	215	214	221	224	225
195	200	204	207	209	211	212	212	210	211	218	218	214	219	223	225
202	207	211	214	215	216	217	219	220	220	225	222	215	217	222	226
213	219	224	226	226	226	225	230	235	234	233	226	215	214	220	228
225	232	238	241	239	238	238	244	252	250	243	232	217	211	220	232
239	245	254	257	254	252	251	259	271	267	254	238	220	209	219	237
250	257	267	271	268	265	265	275	288	283	265	246	225	210	221	242
254	259	268	274	275	277	280	290	298	293	279	265	244	227	234	254
258	258	268	276	282	288	296	305	307	303	294	284	265	246	248	266
261	258	268	278	288	299	311	320	317	314	309	304	287	264	263	280

(a) elevation

1	5	1	3	3	3	3	3	3	1	1	3	3	3	11	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	1
1	1	3	1	1	1	1	1	1	1	1	1	1	1	1	1
1	2	6	7	2	2	2	1	1	2	2	4	1	1	1	1
1	2	16	5	3	3	2	1	1	2	3	8	2	2	2	1
1	20	6	1	4	3	1	1	2	2	3	14	3	3	2	1
1	28	2	1	4	1	1	2	2	3	20	4	4	3	2	1
30	3	2	5	2	1	1	2	3	26	65	19	4	3	2	1
34	3	6	3	2	1	1	2	32	66	2	41	10	3	2	1
229	164	154	147	128	2	1	115	74	5	1	1	40	3	2	1
23	4	3	3	16	124	119	1	6	2	1	1	35	3	2	1
19	3	3	2	2	15	3	2	2	3	1	1	30	3	2	1
15	3	2	2	1	1	14	2	1	1	2	3	4	22	3	1
11	3	2	1	1	2	9	2	1	1	2	3	3	12	1	1
8	2	2	1	1	2	4	2	1	1	2	2	2	5	4	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
5	1	3	3	3	3	3	3	1	1	3	3	3	11	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	4	1	1
1	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1

(b) Accumulation values

Figure 8.1: A sample of elevations and their accumulation values.

At this step, the user needs to input the location of the outlet which will be the starting point for the extraction of stream channels and watersheds. This program will only extract the stream channels and watershed draining to the selected outlet.

4). Creating a stream network

This step delineates stream channels based on the drainage accumulation values. The user needs to input a threshold value as a criterion to define stream channels. Only those cells with an accumulation value exceeding the threshold will be delineated. With this threshold value, the user can extract stream channels of different scale and can accommodate the resolution of the DEM data when extracting channels.

5). Coding stream/segments/finding segments lengths

This step is optional. It will code the stream network extracted from step 4. Each segment of the stream networks will have a unique identifier and the length of each segment can be calculated.

6). Finding sub-watershed basins

This step finds the sub-watershed. It will assign a unique number for each sub-watershed. Therefore, users can further pursue the characteristics of each sub-watershed.

This program will only extract the watershed pertaining to the outlet. Using this program to identify the drainage networks on a whole data set, the user will need to locate the outlet for each watershed and run the program once for each outlet. This will be a tedious and time-consuming process. The requirement of interactive input will hinder the option of running the program in batch mode. One advantage of batch mode execution is that users can run the program in a particular time, e.g., mid-night when the computation load is low, and make better use of computer resources. Also, with batch mode, users can run several jobs at once without interference. To these users, the lack of batch mode execution is a major drawback.

The program in GRASS 4.0 version is called 'r.watershed' and was developed by Ehlschlaeger (1989). Basically this program follows the same procedure as the one in version 3.1, from smoothing, identifying pits, computing accumulation value, the extraction of channels and watersheds, and the coding of results. It also requires threshold values for watersheds. However, the implementation of this program brings in some advantages to this program over the previous version. First, this program will take command line parameters so that user can run it in batch mode without constraint on running time and number of files processed. Secondly, this program will extract all the channels and waterbasins inside the data area in one execution. In addition, this program also provides optional information for further analysis. Upon request, it will calculate S (slope) and LS (slope length) factors of the Revised Universal Soil Loss Equation (RUSLE). It also provide information to interface with ARMSED, a software for modelling storm-water runoff and sedimentation yield. Interested readers should refer to the GRASS 4.0 User Manual (U.S. Army CERL, 1992) for more details.

Several tests have been conducted to test the performance of these two programs. In the following section, this author documented the processes and finding of these tests.

8.3 Experiments

Both versions have been installed in a Sparc Station 2 computer for this current research. A series of tests has been conducted to test the performance of those

extraction functions of the GRASS software. The process and findings of these tests are documented next.

The 'watershed' program of GRASS 3.1 was tested using the Kaneohe DEM. An outlet in the west side of the data set was chose for tests. The resulting channels and watersheds are shown in Figure 8.2 and 8.3.

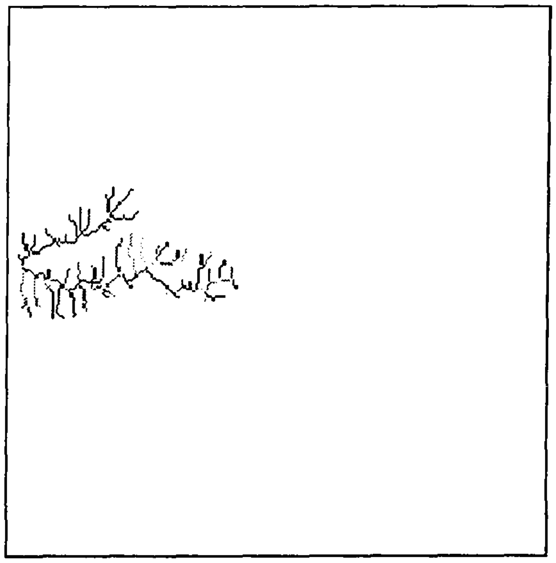


Figure 8.2: Stream channels extracted by GRASS 3.1.

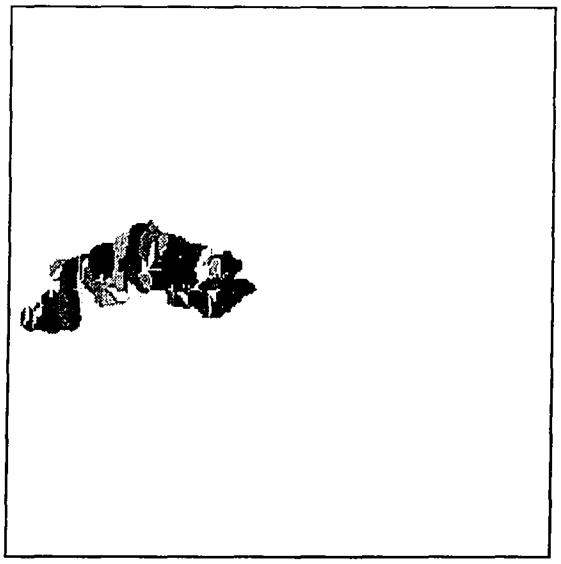


Figure 8.3: Watersheds extracted by GRASS 3.1.

As aforementioned, this program only extracted watersheds and channels embracing the specified outlet. It may be useful for detailed study of a single water basin. For the extraction of channels and ridges in a large area, this function is rather limited.

The ‘r.watershed’ of GRASS 4.0 was used for the same data set. For this test, a threshold value of 500 cell was used for the extraction of watersheds. The resulting watersheds and channels are shown in Figure 8.4.

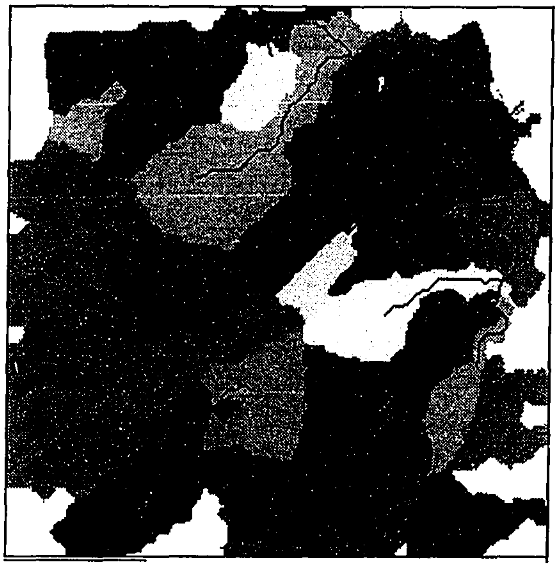


Figure 8.4: Watersheds delineated by GRASS 4.0.

By assigning different threshold values, users can delineate channels and watersheds of different levels. For the extraction of stream channels, the options of various levels can be done with manipulation of the drainage accumulation value. In running 'r.watershed', the program will create a data layer of drainage accumulation. Those cells at the foot of a slope receives water from their upper slopes and have an accumulation value reflecting the slope length. For example, on a DEM with a 30 meter resolution, the cells at the end of a 300 meter wide slope may have an accumulation value of 10. Compared to those cells in the stream channels, the accumulation from upper slope is rather limited. Those cells on the stream channels will have input water from both sides of slopes and from their upper streams, thus have very large accumulation values. In a test, those cells with an accumulation value over 300 were extracted, as in Figure 8.5. Using the same accumulation data, this author also extracted those cells with an accumulation value exceeding 50 (Figure 8.6).

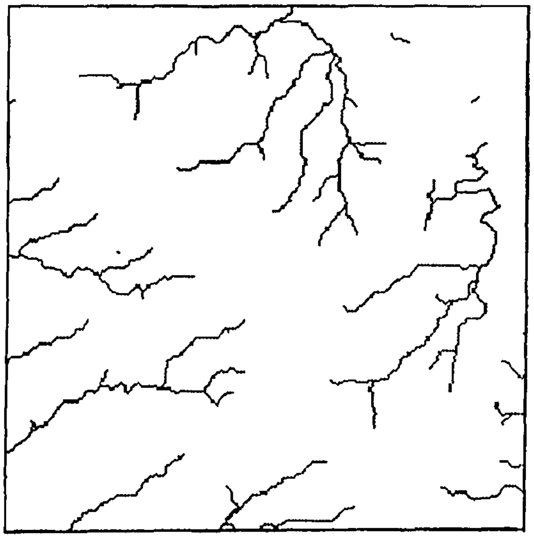


Figure 8.5: Stream channels extracted by a threshold of 300 cells.

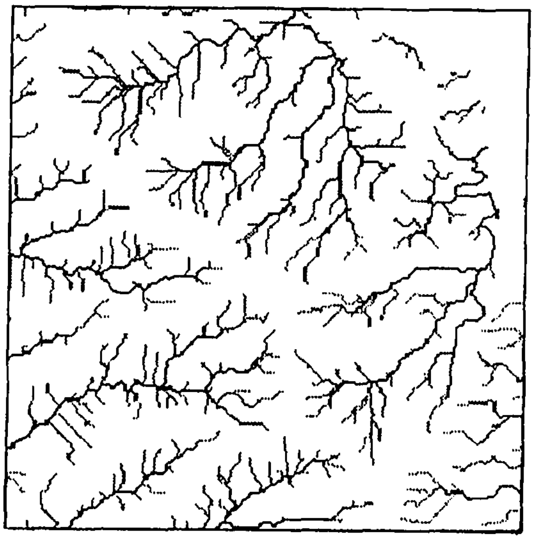


Figure 8.6: Stream channels extracted by a threshold of 50 cell.

By viewing the results so far, the performance of these programs is fairly good. The stream networks are well connected. The locations of channels fit the terrain relief very well when checked with locations and their elevations. However, these programs only extract stream channels, and not ridges.

Several experiments have been conducted to extract ridges using GRASS. The first test was based on an intuitive idea that ridges are divides of overland flow, therefore, they are the cells without any input water from their neighboring areas. This author retrieved the accumulation data calculated from previous steps and extracted all those cells with an accumulation value of one. It was expected that those cells would indicate the network of ridges. The result of this test is shown at Figure 8.7.

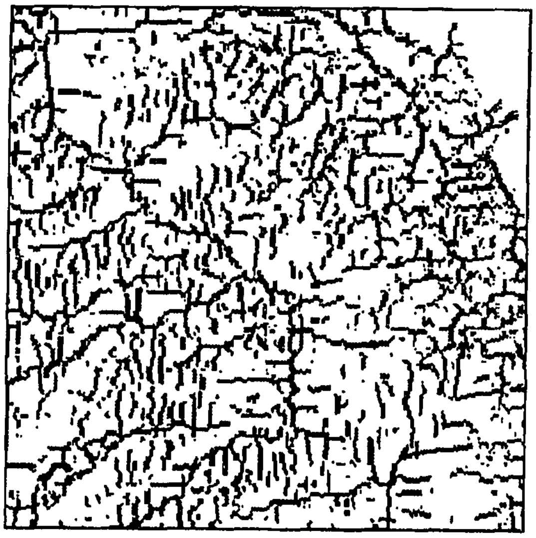


Figure 8.7: Points without water input from their neighbors.

As the figure shows, the ridges from this test are not well connected. The result is similar to those results extracted from profiling methods tested in the previous chapter. This similarity is understandable. Some 'ridge' cells are located on saddle areas and receive water flow from nearby ridges, therefore their accumulation value is larger than 1. This explains why the major ridges are not well connected. On the other hand, there are many minor local maximum with an elevation higher than its neighbors, yet their scales are too small to be recognized as ridges. Basically, this approach is similar to the local profiling method tested before. It will carry the same disadvantages and limitation as the local profiling methods, which have been investigated before. This current test does not merit further exploration.

Although the 'r.watershed' function of GRASS 4.0 does not delineate ridges explicitly, it does delineate water basins and, consequently, the boundaries of watersheds. There is an overlay between ridges and the divides of water basins. Ridges, by nature, are the divides of water flows. There are exceptional cases for ridges not dividing watersheds, yet in general, ridges form the boundaries of watersheds, this is specially common in mountain areas. By delineating the watershed boundaries, one can delineate major ridge lines. To delineate the boundary of watersheds, the 'r.poly' function of GRASS 4.0 was used to extract the boundaries of watersheds, as shown in Figure 8.8.

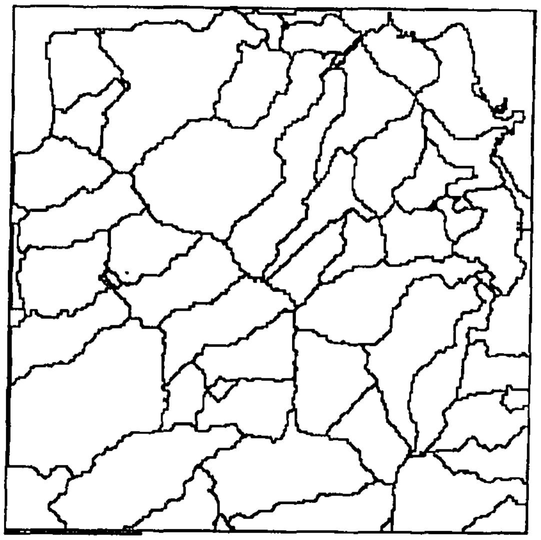


Figure 8.8: Boundaries of watersheds.

The watersheds boundaries have a large intersection with ‘ridges’, yet there is much discrepancy as well. The network of watersheds does not include the tributaries that a ridge network normally have. In the real world terrain, ridge networks usually present a dendritic pattern with major ridges and their branches. On the other hand, the watershed network includes divides in plain areas, where ridges are not expected to exist. Thus, this watershed network does not fit our conceptual models of ‘ridges’. To extract ridge lines more reasonably, another test was undertaken.

The following test was inspired by the idea that ridges are the dual of stream channels. They are similar in some ways. For example, on an aerial photo, it is difficult to tell a channel from a ridge by only looking at the tonal contrast. The major difference is that ridges pose a convex shape and channels are concave. But, then if one take a mould of the ground, then ridges on the ground will become channels on the mould. In this test, a false elevation data set was created by subtracting the true elevation of each cell from 1000. This data set is a mould of the terrain surface. The 'r.watershed' function extracted the 'channels' on the new elevation data, which in fact are ridges on the ground. The main interest from this program is the accumulation data. A large accumulation value indicates a major ridge on the ground. Following the same procedures of channels extraction, those cells with an accumulation value over 50 were extracted and categorized. The resulting ridges are shown as at Figure 8.9.

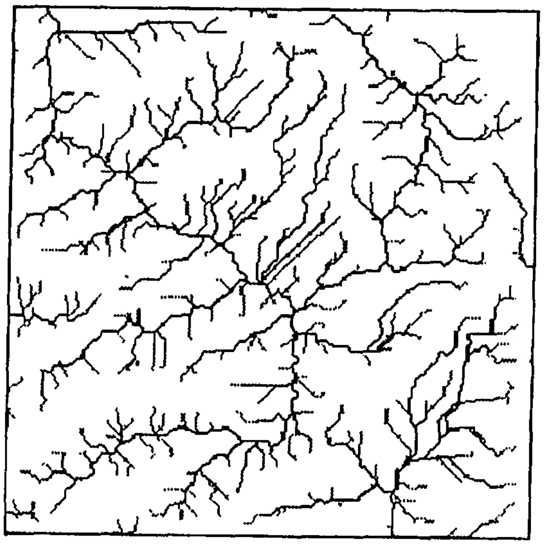


Figure 8.9: Ridges extracted with accumulation values over 50.

The general pattern of this result is reasonable and close to human extraction from contour maps, including ridges of different scales and details. One major advantage of this method is its flexibility for users to define the scale. Using the same accumulation map while with a larger threshold of 300, the extracted ridges are shown in Figure 8.10.

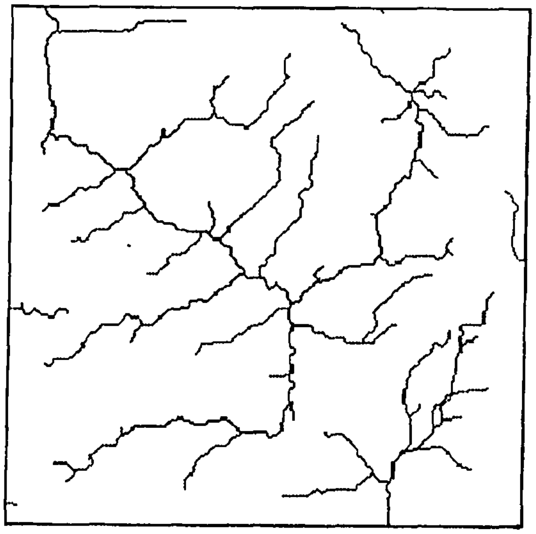


Figure 8.10: Ridges extracted with accumulation values over 300.

At this stage, the purpose of extraction has been fulfilled by the GRASS 4.0 systems using a hydrological approach. The resulting features are continuous and correspond well to the ground surface. The tests of this current study adjourn at this stage.

8.4 Discussion and Conclusion

Although, the implementation of algorithms is not a major concern in this research, a comparison between the two versions of GRASS provide a good example to show the importance of implementation. Both 3.1 and 4.0 versions adopt a hydrological approach, yet 3.1 version requires interactive input from the users and only extract one watershed and the drainage network inside it. In contrast, the 4.0 version is more flexible in execution and extracts all the designated features in one run. The advantage of one over the other is apparent.

Through analysis, it is found that the definition of channels and ridges implied by the hydrological approach reflects human conceptual models of ‘channel’ and ‘ridge’, i.e., they divide slope units and mark the boundary of slope units. Experts do not specify the scale of channels and ridges, yet they can delineate channels and ridges flexibly, although the criteria used are ambiguous and arbitrary. In the hydrological approach, the channels and ridges can be extracted flexibly and clearly. The criteria, or the threshold value, can be defined clearly and applied to various data sets. One of the original appeals of the development of automatic extraction is their objectivity. The hydrological approach fulfills this goal.

The result shows that features extracted from the GRASS software are well connected. The noise and local minor relief for both ridges and channels do not show up on the final result. The reason is that the hydrological approach is not based on local relief, but rather based on flows of a large area. Also, the threshold value screens out those candidates for channels and ridges which are not significant enough to be extracted. The combination of these factors take care of the noise in DEM data and the true minor relief on the ground.

The hydrological approach requires intensive computations. These computation are next to impossible for humans to conduct, yet they are suitable and simple enough for computers to handle. Each of the tests conducted here took less than 10 minutes on a Sparc Station II machine. In all factors of concern, the performance of hydrological approach in general, and the GRASS software in particular, can meet the requirement of various research needs in providing consistent while flexible extraction of channels and ridges.

Chapter 9. Comparisons and Discussion

9.1 Comparisons

The four proceeding chapters have tested the performance of four groups of automatic methods of extracting ridge and valley lines. The investigation on the extraction of terrain features from TINs concluded that the feasibility of such an extraction method relies on the generation of TINs. A series of tests were conducted to reveal problems arising from the generation of TINs from grid DEMs, one of the most common sources for generating TINs. The resulting TINs differed in the distribution of nodes and the connection of triangles. They were created purely from the geometric properties of the nodes and triangles, without referring to ground truth. At the best, they represent a crude approximation of the ground surface. After conducting a series of experiments to generate TINs by various methods, Kumler (1992) concluded that the TINs resulting from regular-grid DEMs and digital contours leave much to be desired in terms of accuracy and storage efficiency. This conclusion reinforces this author's belief that the construction of a satisfactory TIN needs further improvement. Without reliable TINs, the extraction of terrain features from TINs is not meaningful. In fact rather than extracting ridge and valley lines from TINs, many researchers have suggested that valley and ridge lines could be used as a basis for the construction of more reliable and accurate TINs (Douglas, 1986). A further extension from such thinking is that, rather than being extracted from a TIN structure, ridge and valley lines should be extracted first as skeletons that are to be used in the construction of a TIN. This author believes such a methodology is reasonable and proposes a new way to integrate the valley and ridge lines into the process.

Several extraction methods pertaining to the three groups of profiling, tracing, and hydrological approaches have been investigated through a series of tests. These methods are compared in this chapter. The performance of various methods within each of the three groups are generally similar. Therefore, only one method was chosen from each group for further evaluation: the discrete profiling method from the profiling group, the streaming algorithm from the tracing group, and the 'r.watershed' command of GRASS version 4.0 from the hydrological group. Since the same extraction methods are applicable to the extraction of both valley and ridge lines, for the sake of conciseness, the discussion here will deal with ridge lines only.

The evaluation of these methods starts with a visual inspection. Three figures have been prepared schematically to show their performance. Figure 9.1 through 9.3 all contain a hypsometric map overlayed with hill-shading that serves as a basemap. The basemap shows the ground relief so that viewers can envision the locations of ridges and valley lines. On top of the basemap, terrain features resulting from each of the three selected methods are plotted. Viewers are advised to examine the amount, location, and continuity of the extracted features and the difference of spatial patterns between the three figures. These qualities will be used to compare these methods in following paragraphs. Figure 9.1 shows the result obtained from the profiling method; Figure 9.2 shows that of the tracing algorithm; Figure 9.3 shows that from the hydrological approach of GRASS software.

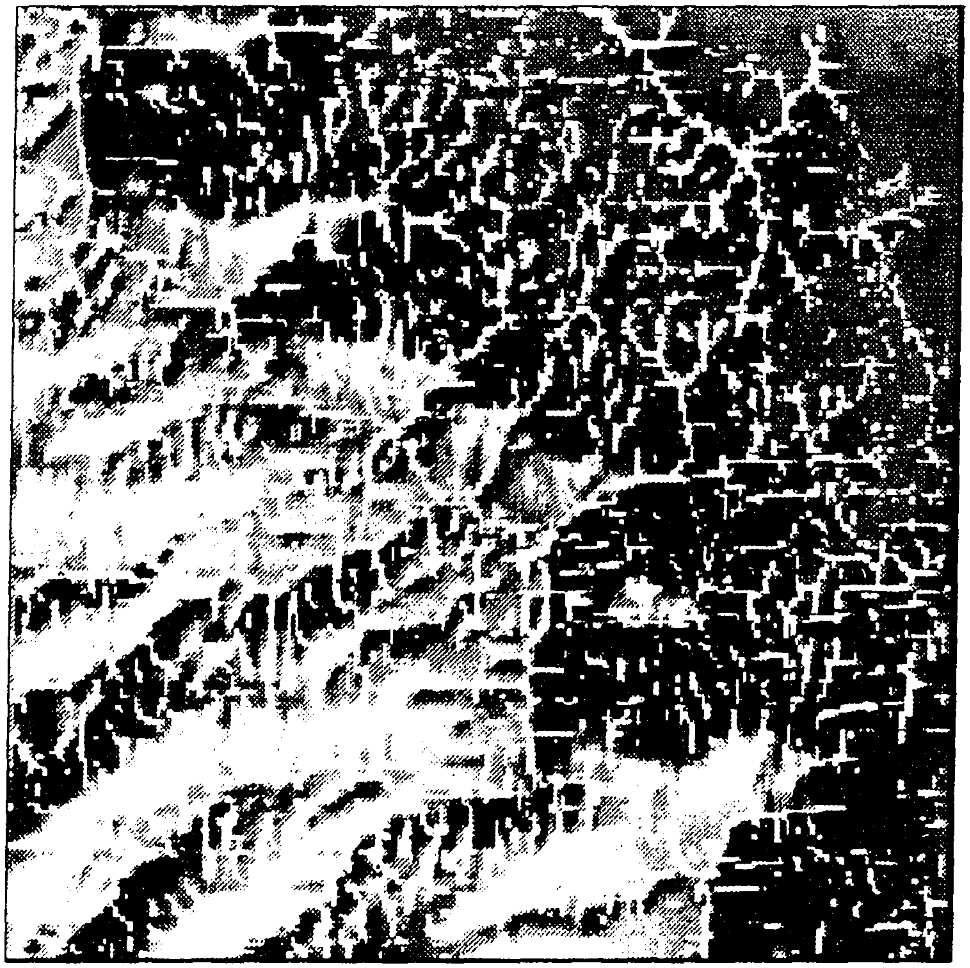


Figure 9.1: Ridges resulting from a profiling method.

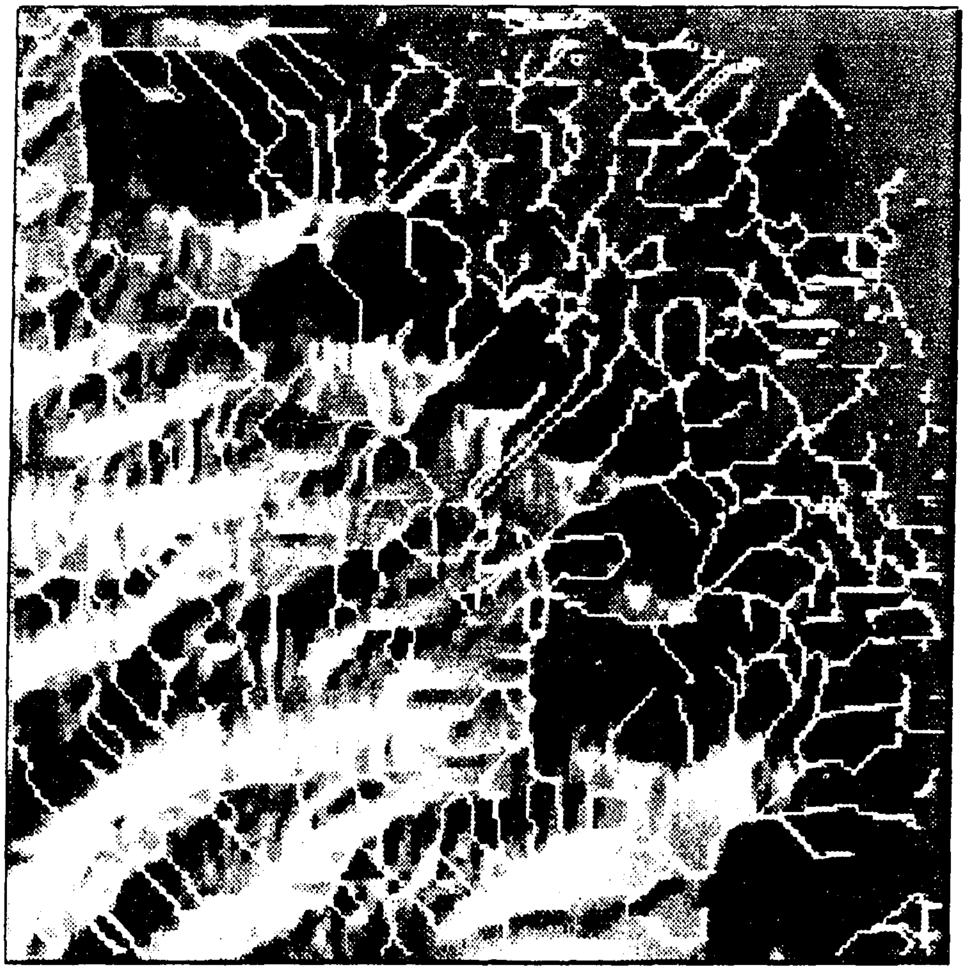


Figure 9.2: Ridges resulting from a tracing method.

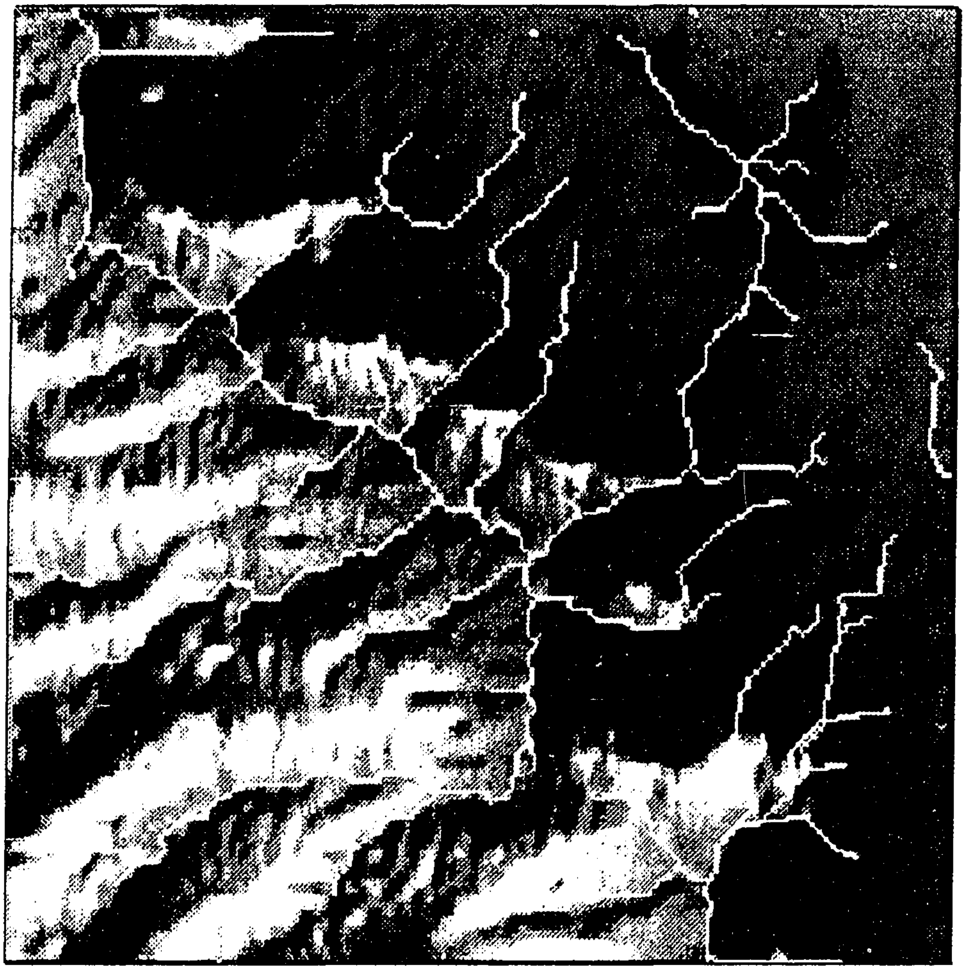


Figure 9.3: Ridges resulting from a hydrological method.

Although these figures may effectively reveal the performance of each method visually, a parametric comparison is in order. In order to compare these methods more comprehensively, the extracting results will be examined in four aspects. They are: the number of extracted features, the continuity of features, the positional accuracy of features, and the agreement among the three tested methods. Details of these examinations follow.

9.1.1 Number of Extracted Features

Before further discussion, it should be mentioned first that the resulting ridge lines are in fact composed of individual pixels. The set of these pixels represents the ridge network. As listed in Table 9.1, the profiling approach extracted 11,876 pixels, the tracing algorithm extracted 6,692 points, and the function of GRASS 4.0 extracted 4,854 pixels. These numbers of pixels respectively represent 18.12 percent, 10.21 percent, and 7.41 percent of the pixels of the test data set respectively. The percentages are somewhat high compared to a manual interpretation will normally extract. Many of these pixels clump together in small areas and thus do not reflect the length of ridge lines. To make the pixels more meaningful in representing the extension of the ridge network, a line-thinning function was adopted in GRASS to trim off those pixels adding extra width to the ridge lines. This process effectively reduces the amount of pixels in each result. The numbers of pixels removed by the line-thinning process are listed in Table 9.1.

Table 9.1: Numbers of pixels before and after the line thinning process.

	Profiling	Tracing	Hydrological
Original No.	11876 (18.21%)	6692 (10.21%)	4854 (7.41%)
No. after thinning	8751 (13.35%)	6075 (9.27%)	4598 (7.02%)

All ridge lines are of single-pixel width after thinning, therefore, the numbers of pixels are an approximate measure of the total length of ridge networks. As a result, the profiling method extracted 13.35 percent of the test data as ridge lines, followed by 9.27 percent of the tracing approach, and 7.02 percent by the hydrological approach. The profiling method extracted the most points that spread over almost all parts of the area (Figure 9.1). The ranking of the amount of resulting features is not the focus of this discussion, however. What is of more concern in this comparison is the flexibility of each method in extracting features of various levels of hierarchy.

Users may desire extraction of different numbers of features depending on the use to which the extracted data will be put. The hydrological approach readily provides such great flexibility. Within GRASS, users can select the accumulation value as a criterion to extract terrain features in various detail. As tested in Section 8.3, this author used GRASS to extract two sets of ridge lines, one with an accumulation value of 50 and the other with 300. The former yielded 4854 while the latter 2074 pixels. Figure 9.4 shows the ridge lines in dark that are dropped out by raising the criterion.

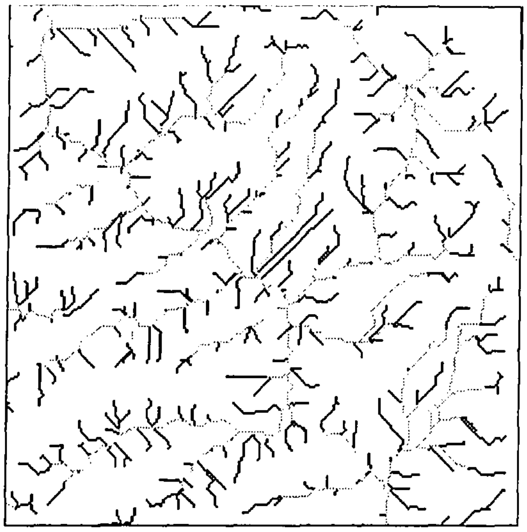


Figure 9.4: Ridge lines (dark) dropped out by adopting a higher accumulation value.

The ridge network extracted by the accumulation value 300 is a subset of that extracted by the value 50. Both networks show a high degree of connection. Their difference appears to be the number of branches and extension of ridge networks. These two results demonstrate the flexibility of the hydrological approach in extracting ridges. The accumulation value can effectively categorize ridge and valley lines into different levels so that users can extract various levels of features as desired.

The tracing method does not provide flexibility in extraction. As addressed in Section 7.3, there exists ambiguity in defining the starting points for tracing. Yet, once a criterion is selected, the number of extracted starting points will be fixed. The nature of the tracing process itself imposes further rigidity because that no hierarchical extraction pertains to the tracing process. As a result, the tracing approach provides no flexibility in extracting various numbers of points.

The profiling method adopts a clear-cut rule in extracting ridge lines. It appears to be non-flexible in defining features. However, as the extraction is performed by a local processor, there exists possibility to extract various numbers of points by changing the criteria adopted by the local processor. A test in Section 6.2 was conducted to test such a possibility: the criterion for defining a ridge point was raised from one meter higher than its two neighbors to two meters higher, resulting in a reduction in number of extracted points from 7626 to 5906. The points dropped out by the stricter criterion are shown as dark points on Figure 9.5. The shaded points are the remaining points after raising the criterion of profiling approach to two meters.

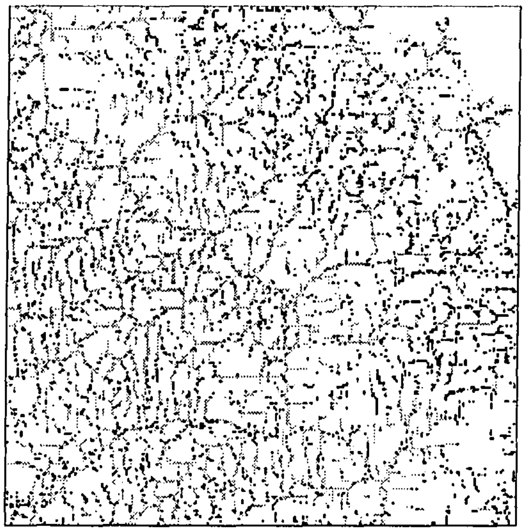


Figure 9.5: Ridge points (dark) dropped out by the 2 meter criterion.

As shown in the figure, the variation in criterion does not change the extension of the ridge networks. Instead, the lost data appear to be points scattered all over the area. The effect of stricter criterion on the profiling method reduces the numbers of short and noisy features, but it also breaks the major ridge lines more frequently compared to a looser criterion. Variation in criterion does not facilitate a flexible and hierarchical extraction of features in this test.

In summary, the profiling method extracts the greatest number of ridge points, followed by the tracing method and the hydrological method. Yet, such a ranking does not tell the whole story. The ability of each method to extract ridge lines at various hierarchy is a more important concern. Such flexibility is present in the hydrological method only.

9.1.2 Agreement Analysis

The agreement analysis reckons the numbers of same points extracted by different combination of methods. The purpose of this analysis is to examine correspondence between the resulting ridges. The results of the analysis are tabulated in Table 9.2. All layers presented here are the original results without thinning.

Table 9.2: Cross-table for numbers of points extracted by different combination of methods.

	H	T	H&T	NH&NT
P	1619	2105	1813	6339
NP	1137	2489	285	49749

note: H: Points extracted by hydro-method.
T: Points extracted by tracing-method.
P: Points extracted by profiling-method.
NH: Points NOT extracted by hydro-method.
NT: Points NOT extracted by tracing-method.
NP: Points NOT extracted by profiling-method.
H&T: Points extracted by both tracing and hydro-method.
NH&NT: Points extracted by neither tracing nor hydro-method.

Of the 65536 points in the whole area, 49749 points are not extracted by any of the three methods, and 1813 points are extracted as ridge points by all three methods. Obviously, the profiling method extracted many more points (6399) that are not extracted by the other two. In contrast, the number of points extracted by both of the other two methods while not by the profiling method is only 285.

In a pair-wise comparison, 70.70 percent of the points extracted by the hydrological approach are also extracted by the profiling approach whereas only 43 percent of points extracted the hydrological method are also extracted by the tracing approach. Among those points extracted by the tracing approach, 58.55 percent of them are also extracted by the profiling method, while only 31 percent of them are also extracted by the hydrological approach. These figures show that among the three pairs of methods, the profiling and hydrological methods have the closest fit in extracting ridge points.

These statistics only represent the agreement on three particular layers of resulting ridges. To confirm the tentative conclusion on the agreement of the three methods, the agreement on the other three sets of ridges is examined. These three layers include: the original points extracted by the profiling method, the thinned layer of the points by the tracing approach, and the points by the hydrological method with accumulation value 300. Numbers of points extracted by each combination of methods are displayed in Table 9.3.

Table 9.3: Cross-table for numbers of points extracted by different combination of methods (thinned data sets).

	H	T	H&T	NH&NT
	=	=	====	======
P	1521	3918	967	7404
NP	553	2652	122	60485

Based on this set of results, 73.34 percent of points extracted by the hydrological approach are also extracted by the profiling method and 58.55 percent of points extracted by the tracing approach are also extracted by the profiling method too. Only 52.51 percent of points extracted by the hydrological method are extracted by the tracing approach. The statistics of both sets of data demonstrates that the profiling method and the hydrological method have a best match among the three pairs of methods.

9.1.3 Continuity Analysis

Ridge lines of a mountain area are normally continuous. This part of analysis investigates the continuity of ridges resulting from the three methods. This study counts the numbers of separate disconnected ridge networks on the ridges extracted by each method. In this analysis a "network" is simply defined as a system of ridge lines separate from other system of ridge lines. The number of extracted points is divided by the number of networks, gives an average length per network. The number of networks and the average length of networks indicate the continuity of each layer. The implementation of this analysis adopts some theorems of graph theory.

A ridge network can be thought of as a *tree*, a connected graph of nodes and edges without circuits (Johnsonbaugh, 1984, p.159). The theorem that a tree of n nodes contains n-1 edges can be proved through induction. Figure 9.6a is a schematic demonstration of the theorem: the first edge of a tree requires two nodes; thereafter, the number of edges added to the tree will introduce the same number of new nodes. For a graph that contains circuits and thus forms faces (areas), a more general theorem states that on a connected graph, the number of nodes plus the number of faces equals the number of edges plus one, as shown in Figure 9.6b.

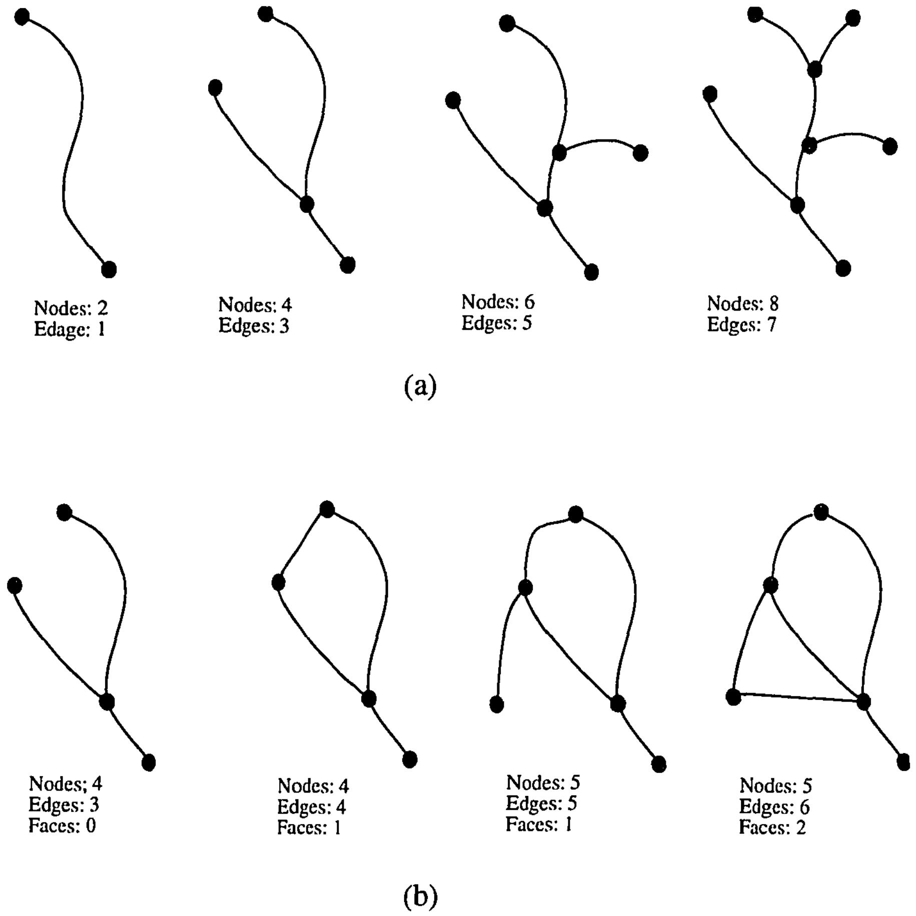


Figure 9.6: Nodes and edges on a tree.

The relation between the numbers of nodes, edges, and faces can further derive the number of separate graphs in a data set. On a data set that contains separate graph, the numbers of nodes and faces minus the number of edges will indicate the number of separate graphs. In the data set here, each connected graph represents a separate network of ridge lines. Here the GRASS software is used to convert the three layers of ridge points into three separate vector files. Before the raster-to-vector conversion, the ridge layers need to be thinned to single-pixel width. The numbers of nodes, edges, and areas, as well as the numbers of networks and average number of points on networks are also obtained from the vector files. Table 9.4 lists these numbers for three layers resulted from: the profiling method with a 2 meter criterion, the tracing method, the hydrological method with an accumulation value 50.

Table 9.4: Number of nodes, edges, and faces on each resulting ridge map.

	No. of Nodes	No. of Edges	No. of Networks	No. of Points	Average No. of points per Network
Profiling	2272	1635	649	5906	9
Tracing	1139	953	251	6075	24
Hydrological	831	873	11	4598	400

It is recognized that the thinning process of GRASS introduces many artificial loops when lines are close together. Nevertheless, the existence of such loops does not affect the number of networks and can be ignored for the purpose here because of the concern here is the number of networks.

The result obtained through the hydrological method contains only 11 separate networks, the smallest number among the three results. Its average number of points per network is about 400, the largest among the three. Both numbers of networks and average points indicate that the result obtained from the hydrological method is best connected. On the negative side, the main ridge network of the test area was extracted as two separate networks on the hydrological output. The disconnection appears on a major saddle area. This disconnection reveals the weakness of the hydrological approach in certain terrain types. Points around saddle areas are less likely to have a great accumulation value because they are part of ridges as well. Thus, they are often missed in the extraction process. Except for the possible disconnection, the hydrological method generally results in well-connected networks.

The other two methods introduce many more networks than a manual interpretation will normally extract. The profiling method extracted 649 separate networks and the tracing method extracted 251 separate networks. The average number of points for networks is 9 for the result obtained from profiling method and 24 for that from the tracing approach. The numbers in Table 9.4 show that the hydrological method delineates ridge networks that are mostly continuous, followed by the tracing method. The result obtained through the profiling method is the least continuous.

9.1.4 Positional Accuracy

The evaluation of positional accuracy of these methods is subtle. Positional accuracy refers to the difference between the location of extracted ridges and the true location of ridges appears on the DEM data. The subtlety arises from the fact that there is no standard answer for extracting ridge and valley lines either conceptually or technically, as discussed in Chapter 1 and 4. Yet, a standard is needed to evaluate the accuracy of ridge locations. To resolve the problem, a compromise has been made by using manually extracted ridge lines from a contour map as the standard for later evaluation. A contour map was generated from the gridded DEM data using ARC/INFO software, instead of using an existing USGS topographic map. Such practice is to avoid the errors may arise from the difference between USGS topographic maps and grid DEMs. Some deviation may arise from the grid-to-contour conversion. However, such differences will be less than the resolution of the grid DTM, 30 meter of ground distance in this case. Errors of this magnitude are negligible compared to errors introduced by other possible sources such as manual extraction and digitizing. To minimize the subjectivity of manual extraction and avoid too much detail, only the major ridge lines are manually delineated and digitized. The contour map and the manually extracted ridges are shown in Figure 9.7.

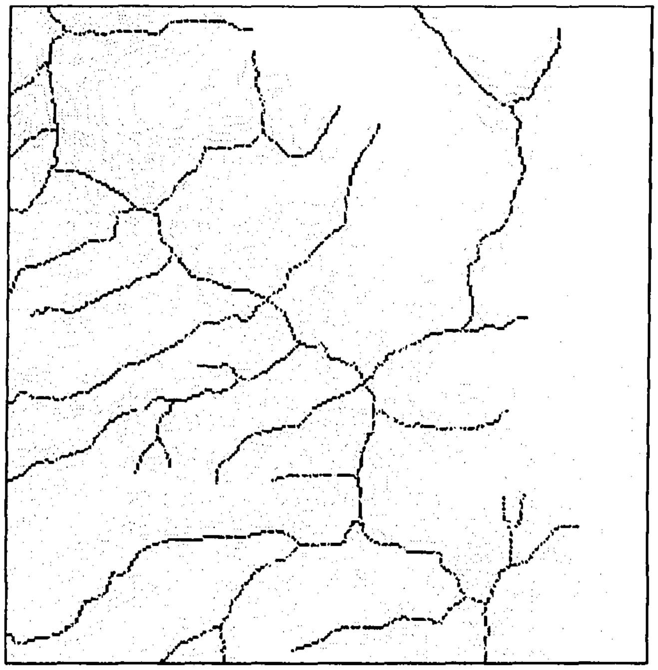


Figure 9.7: Contours and manually extracted ridges.

A buffer map for each of the three layers of resulting ridges has been created by the 'r.buffer' function of GRASS. These buffer maps display the distance between a point and its nearest non-zero point, i.e., the distance between each point and the nearest ridge point. Then, the map of the standard ridge lines is overlaid separately with the three buffer maps. The results of the overlaying show the deviation between the standard ridge lines and the corresponding ones extracted by each method. The values of deviation are presented in Table 9.5.

Table 9.5: Positional accuracy of three methods.

distance	Profiling	Tracing	Hydrological
0	643	576	707
1	861	739	742
2	255	261	242
3	105	156	112
4	16	51	40
5	4	39	24
6	1	21	6
7	-	18	5
8	-	10	2
9	-	5	2
10	-	6	1
11	-	1	1
12	-	1	-
Total:	1884	1884	1884

Given the arbitrariness of manual extraction, the error introduced by the grid-to-contour conversion, and the errors occurring during the digitizing process, the accuracy of the standard ridges itself is questionable. Taking all these errors into account, it is proposed here that deviation within two-pixel width is a reasonable tolerance value. Namely, the extracted ridges within two pixels of the standard ridge will be deemed correct. Based on such a criterion, table 9.5 shows that 93.36 percent of the standard ridge points fit with the results of the profiling method. The same table shows a 83.65 percent fit with the result by the tracing algorithm and a 89.76 percent fit with the hydrological method. In conclusion, locations of ridge points extracted by the profiling method is most close to the true locations of ridges, followed by those obtained through the hydrological approach and the tracing approach.

9.1.5 Summary of Comparisons

The above analyses address the advantage and disadvantage of the three methods from various aspects. The properties are summarized in Table 9.6.

Table 9.6: Summary on the properties of three methods.

	Amount of feature	Continuity	Positional accuracy
Profiling method	Not flexible.	Least continuous.	Fit best with true locations.
Tracing method	Least flexible	Not continuous.	Least fit with true locations.
Hydrological method	Very flexible, can distinguish feature hierarchically.	Well connected except in certain areas such as saddle areas.	Mostly fit with true locations

Obviously, the hydrological approach performs the best in the aspects of continuity and flexibility. Although the hydrological method is ranked second in positional accuracy, its results are compatible with the standard ridge up to 87 percent, a figure that may be good enough for many applications. It can also be noticed that the profiling methods fit most closely with the standard ridge lines. The profiling method also shows a high agreement with the other two methods. Yet, the fragmentation with this method is severe, which may be a serious drawback for many applications. The tracing method ranks low in all three aspects of concern. As a result, this author deems the hydrological approach to be the most satisfactory method for general-purpose extraction of ridge and valley lines.

The above summarized difference can be attributed to several factors in different perspectives. The following section discusses the relationship between the performance and those factors.

9.2 Discussion

The automatic extraction of terrain features is relevant to several factors: nature of DEMs, human cognition, and implementation to computers. From a conceptual perspective, the extraction process categorizes the ground surface reflecting human definitions. Therefore, the evaluation of these methods must refer to human conceptual model of terrain features. These methods attempt to duplicate manual operations in the computer algorithms. The nature of terrain and the characteristics of DEMs are reflected in the performance of extraction methods. The performance of each method is an integrated result of how it interacts with these factors. This section analyzes the performance of the three methods from these multiple perspectives.

In the profiling method, a ridge line is defined as a set of points carrying a convex cross profile. Subsequently, terrain features can be defined by geometric properties. After all, extraction of terrain features is fundamental to morphometry that describes the geometry of the earth's surface. Besides being used by a variety of profiling methods, geometric properties are also adopted in several tracing approaches for the selection of starting points. From the modeling aspect, not all convex patterns on the ground will be presented on the grid DEMs. Because the U.S. Geological Survey's DEMs record the ground height with integer values, it is likely that some minor relief on the ground will be truncated and will not show in the DEMs. Besides, the discrete sampling of the DEM data may skip a convex profile sitting between two DEM points, as shown in figure 9.8.

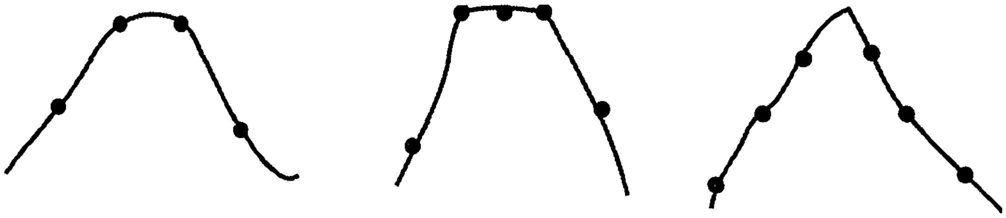


Figure 9.8: Ridge points missed for not having a convex profile.

The discrete nature of DEMs raises another intricate question: is it sensible to measure shape at a fixed scale? To answer this question, one might think of the fractal geometry theory that states that many geometric patterns of terrain features are scale-independent and can be observed at various scales (Mandelbrot, 1983). However, such independence does not hold on a discrete data model such as a DEM since those patterns smaller than twice the data resolution will be lost. Therefore, in theory, the identification of the terrain features is closely affected by the scale one adopts. Whether a cross profile is convex or not depends on the scale of the representation. Using a fixed-size filter to examine the shape of cross profiles will surely miss some candidate features. For instance, Figure 9.9 shows a portion of the test DEM data where the main ridge is disconnected.

```
642 633 622 607 590 577 564 551 539

648 639 630 617 602 592 581 572 564

640 633 627 616 606 599 594 589 587

627 623 620-612-606 605 605 606 610
         |   |   |
614 613 612-607-607 611 616 623 633
         |   |   |
605 608 608-604-606 613 621 630 644

605 608 607 603 604 609 616 625 638

605 608 605 602 603 606 611 619 632

603 606 601 599 599 602 607 615 626
```

Figure 9.9: Sample DEM of the test area with main ridge disconnected.

In view of the elevations of the whole area, clearly there exists a saddle area in the central part, which is concave along the direction of a major ridge and concave in the perpendicular direction. Such convex pattern will not present when we evaluate only the elevations of the 3x3 window at the center. This example demonstrates that not every ridge point will pose a convex cross profile given a fixed scale. As a result, some ridge lines are disconnected in several places when the ridge points are missed by the fixed-size filter. A possible remedy to this deficiency may be achieved by comparing the cross profiles at several scales, e.g., 5x5 window, 7x7 windows, etc.

However, the additional tests will extract many more ridge points to a set of points which needs to be trimmed. At scales larger than the data resolution, the self-similar roughness of terrain surface, as indicated by its fractal dimension (Mandelbrot, 1983; Lam *et al.*, 1993), is more likely to show up on the DEM data. It is anticipated that there will be many inflections along the terrain profiles shown on DEMs, therefore, many points are likely to show convex/concave pattern along their profiles. Such abundance of points with convex profiles is shown in figure 9.1. The disconnection of ridge lines and abundance of scattered points suggest that the shape of cross profiles alone is not comprehensive in extracting terrain features. In addition to analyzing the shape of cross profiles, more comprehensive factors should be used to define ridge points. For example, convex points adjacent to major ridge lines at a high elevation should be more likely to be ridge points in comparison to those with the same profile in some low flat area. To make the profiling method useful, identification of such rules is needed.

The tracing approach emulates human delineation process in that a ridge is taken as the top portion of a mountain range. A tracing process starts from a feature point of terrain, e.g., a peak or a saddle point. The identification of a starting point is subject to errors, as points are selected on the basis of geometric pattern shown on DEM data that has a finite resolution as discussed previously. In addition to problems in selecting starting points, the rules adopted by automatic tracing are too simple to be effective. The manual delineation process performed by interpreters involves much heuristic knowledge. Human heuristic knowledge is often vague and *ad hoc* and thus is difficult to specify and implement in algorithms. There are situations in which a simple rule such as climbing along the steepest direction will not result in the correct path as expected. These situations are investigated in the literature of artificial intelligence (AI) as an analog to describe some shortcomings of hill-climbing search (Winston, 1992, p.73). Hill-climbing is a searching method in AI that applies heuristic approach to measure the remaining distance to a goal. Winston (1992, p.73) listed three types of conditions that the hill climbing process will fail. 1). The *foothill problem* crops up when more than one peak sits in a nearby area, as illustrated in figure 9.10a. The trail with a maximum immediate gain in elevation may lead to a local maximum but miss the global maximum. An implication of such problem to the tracing process of terrain features is that computers fail to identify the major ridge while tracing along a trivial one. 2). The *plateau problem* occurs when the tracing process confronts a flat area with separating peaks, as shown in the example in figure 9.10b. No obvious direction can be taken for further tracing; eventually the tracing will be terminated. 3). In the *ridge problem*, the search process reaches a false peak which in fact is only a point on a ridge, as in the example of figure 9.10c, caused by the limited number of search directions and step sizes. The limitation of searching directions here also implies limitation of possible solutions in AI. For tracing on a DEM, the limited number of neighboring points poses serious difficulty for the operation. Once it derails, the tracing process will follow a false path.

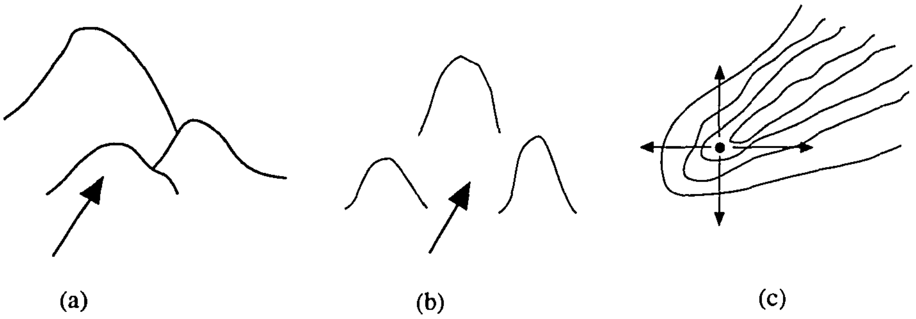


Figure 9.10: Three problems facing the hill climbing process.

When confronting these kind of problems, humans and computers respond differently. These problems are trivial to the human interpreter if he/she can extend the search area far enough to search globally. In fact, that is what human interpreters do during tracing a ridge line. Unconstrained to a local area, they can look farther away. In contrast, tracing methods currently implemented on computers adopt a step-by-step approach. A global search by such approach will be unlikely since the number of searches will increase explosively with the number of search steps taken. To make the tracing process effective, a remedy to the tracing strategy is needed to make them search somewhat globally. Such a remedy may be a major challenge to the software engineering as a whole, though, due to the nature of computer functioning.

The two aforementioned methods attempt to delineate ridge and valley lines in a way that is similar to human interpretation. Their insufficiency reveals a problem in formalizing and implementing human knowledge/behavior into a computer. Humans are strong in symbolic reasoning and pattern recognition while computers are strong in repetitive numerical operations. Direct replication of human knowledge in computers is often not feasible. From the failure of the tracing approach for the extraction of ridge lines, it is concluded that the obstacle to such attempts is in formalizing human heuristic knowledge, which is often *ad hoc* with many rules-of-thumb. Human experts may not be aware of such knowledge when applying it, let alone converting it to rules that can be implemented by computers.

Instead of replicating human process, new approaches should be developed to make use of the computational strength of computers. As cited by Clarke (1990, p.2), Morrison (1980) stated that there were three stages in the adaption of a new technology: 1) the *reluctance stage*; 2) the *replication stage*; and 3) the *full implementation stage*. In the first stage, people are hesitant to adapt the new technology, probably due to the inertia to stay with old technology and the uncertainty of the new technology. In the second stage, new technology attempts to replicate the previous technology. In the third stage, new processes and operations that previously were impossible are emerging. If the profiling methods and tracing approaches stand for research in the replication stage, the hydrological approach can be considered as a full implementation stage of a new technology.

In contrast to the profiling and tracing methods, the hydrological approach avoids emulating human processes by adopting an innovative definition that takes advantage of the computational power of computers. The embedded definition of valley and ridge lines in hydrological approach is not explicit. However, using the accumulation value as the criterion, the stream channels extracted by the hydrological approach are composed of points that are either: 1) located at the intersection of opposite slope complexes having accumulative input from both sides of slope, or 2) the lowest neighbor of an extracted channel point whose large accumulation is inherited by this lowest neighbor. Such a definition of the stream channel flexibly integrates the properties of 'convex shape' and 'the edges of opposite slope complexes'. A selected critical accumulation value filters out points with local minor relief and extracts only those points that are clear enough to be a feature point. It also allows extraction of terrain features on various level of details. Such an integration combined with flexibility of the method explains the success of the hydrological approach in extracting stream channels and ridges.

Nevertheless, the hydrological approach has its own shortcomings. In a flat area where the flow directions of each point are all the same, the only current solution available is to take an arbitrary direction. This situation can be observed in the flat area where the extracting stream channels appear to be straight lines at a fixed angle.

This issue brings up the conceptual issue and *ad hoc* nature of defining a feature. In their study on errors in DEMs, Lee et al. (1991) defined stream channels as points with a smooth gradient in a mountain area. The definition may just be adequate for their immediate purpose. As discussed in Chapter Four, definitions are often *ad hoc* for a specific purpose. The deficiency of hydrological methods occurs in flat areas which can be deemed as a limitation in its definition incapable of emulating human's *ad hoc* inference ability for various terrains. Users of the program should be aware of such type of deficiency. By explaining the sources of deficiency and constraints of each method the advantages and disadvantages of each method are made clearer.

Chapter 10. Conclusion

10.1 Summary

The development of automated method to extract terrain features has the practical aim to spare people the tedious manual work of delineating and digitizing but requires that many theoretical issue related to human cognition, the nature of digital terrain model, and computer software technology be addressed. Ridge and valley lines were the terrain features chosen as the subject of this study. Four groups of methods for the extraction of these features have been identified. They are: symbolic approach, tracing approach, profiling approach, and the hydrological approach. The symbolic approach was not tested because problems arose from the generation of TINs, from which the symbolic approach attempted to extract terrain feature but the other three methods have been tested. Their performance were compared on the basis of the amount, continuity, agreement, and position accuracy of their extracted features.

From the theoretical side, the different performance of these methods is discussed with reference to related factors of automatic extraction. The hydrological approach presents a reasonable answer to all questions about the automation of this process. This method defines ridge and valley lines as points with large accumulation values. Consequently, it is less sensitive to local terrain and produces a more continuous and complete result than the other two methods. The algorithms used for the tracing and profiling were developed to emulate human cognitive processes but the outcomes are not satisfactory. It is concluded that the hydrological approach performs the best on average. The various performances of the three methods present a notion that direct replication of human knowledge into computers may not be feasible in the development of automatic methods. A practical method should be well-adjusted to the functionality of computers and the nature of digital data involved.

This study achieves two purposes. First, it facilitates a better understanding of the nature of different extraction methods so that users of these automatic methods can choose the right tool for a sound result. Second, it addresses the problems facing the development of automatic methods. Future developers of automatic methods can take lessons from this study and adopt a better strategy for methodological development.

10.2 Problem

This research focuses on the methodological aspects of automatic extraction. To remain focused of the subject in the study, the accuracy of DEM data was assumed. However, in reality, DEM data often contains various types of errors. Carter (1989) and Theobald (1989) identified some common errors of DEMs: spikes, strips, and area errors. During this research, the test data of Kaneohe DEM was found to contain strange patterns which strongly suggest errors. In the coastal area, many points with a ten meter elevation are scattered in a neighborhood of zero elevation. To the right of the central part, some straight strips are visible in the hill-shading map. As another example, solid evidence of errors was found in the Diamond Head area of the Honolulu DEM which is adjacent to the Kaneohe DEM. In that case, the elevation in the area is systematically lower than the true elevation. Errors of DEMs were clearly present.

As most suspected errors found in the test data occurred in flat and edge areas, the suspected errors would have not apparent impact on the results obtained by the hydrological approach. Such errors would be critical if located on the main valley or ridge lines. The suspected errors may contribute to the scattered presence of short ridges found in the results from profiling method, since spikes and strips will pose a convex profile in a 3x3 window of DEMs. The same errors would also affect results obtained by the tracing approach as well since many local peaks will be introduced and thus increase the number of starting points for tracing. Users of digital terrain data should check the quality of DEMs.

10.3 Implications

As the results from the extraction process are readily stored in a digital format in computers, many further uses are feasible. In forestry management, where it is necessary to delineate areas near mountain ridges, a buffer zone based on the ridge lines can easily be generated by most GIS packages available today. For practice of soil conservation, the application of universal soil loss equation requires the computation of the slope length as a parameter for modeling the amount of soil eroded. Once ridge lines have been defined, such slope length can be computed efficiently and accurately from DEM data. When properly executed by computers, the extraction of ridge and valley lines will be objective and repeatable and so assure more meaningful applications of terrain features, such as the computation/determination of stream order and drainage density. The automatic extraction of terrain features will promote many more powerful applications of terrain features and facilitate the studies in geomorphology and many related fields.

The concepts built into the automatic methods of feature extraction inspire new concepts in analysis. Many tasks that previously relied on data of ridge and valley lines may take a complete new approach without using those data anymore. For example, the shape of ridge networks that define waterbasins is often subject to hydrological analysis since an elongated waterbasin is more likely to have a longer time lag. At present, prediction before peak discharge following precipitation in watersheds more likely to generate damaging floods could then be determined, such is weak partly because the measurement of shape is crude and partly because the shape of waterbasin is not critical enough to affect the time lag. A major breakthrough on this task is possible by adopting the methodology of hydrological approach of feature extraction. There exists potential for linking accumulation values of points and the travel distance of overland flow between points. If such relation is identified, the amount and the time lag of a flood can be estimated better. This type of application presents an opportunity for future research. The development of automatic methods should not limit itself to attempting replicating manual processes. Instead, it should bring in new powerful tools and methods that are otherwise infeasible without the computational power of modern computers.

10.4 Future Research

The study is a multi-perspective exploration on the automatic extraction of terrain features. The findings present many possible opportunities for further research. There is room for improvement on those methods that are not satisfactory and the author identifies three areas of concentration for future research.

First, the profiling methods for the extraction of ridge and valley lines are currently not satisfactory in several respects, yet they indicate the positions of ridges and valleys most correctly. An improvement in ways to define continuity of feature better while reducing the number of short noisy features defined may make this group of methods useful. One possible resolution is proposed to integrate more parameters in extraction, such as the elevation, the relief, and the length of the extracted features.

Second, the extraction of ridge and valley lines can serve as building block for the extraction of other terrain features. Based on the successful results obtained through the hydrological approach, this author proposes to investigate the extraction of terrain features which have area extent, e.g., valleys and mountains. The idea is partly inspired by the an attempt to integrate the various data bases generated by the U.S. Geological Survey. The Geographic Names Information System (GNIS; USGS, 1983c) produced by the USGS contains the coordinates of a name string of terrain feature, e.g., 'Manoa Valley', but does not indicate the locations of the feature it refers to. If the areal extent of such terrain features can be automatically identified, the various data stored in the GNIS, the DEMs, and the DLGs (USGS, 1983b) may be integrated better and thus foster their applicability.

Third, geomorphology and hydrology probably are the two fields that have most use for the data produced by methods investigated in this study. The concept of accumulation values in the hydrological approach is applicable to many tasks in environmental management and terrain analysis, such as the prediction of flood and modeling of soil erosion. The number of points extracted can be a measure closely related to the complexity of terrain, thus it can be used to derive morphometric parameters and apply to comparative study of different terrain areas.

The development of automatic extraction of terrain features is a multi-faceted task. A successful development will benefit not only geographers but many scientists in related disciplines as well. On the other hand, the development of such tasks requires expertise pertaining to fields in GIS, computer software technology, and many other related fields that apply terrain features for various analyses. Further development on this task is anticipated through the integration of knowledge in related fields.

BIBLIOGRAPHY

Alber, R., J.S. Adam, and P. Gould. 1971. *Spatial Organization: the Geographer's View of the World.* Englewood Cliffs, New Jersey: Prentice-Hall.

Anderson, M.G., and S. Howes 1988. "Computer Simulation in Geomorphology." *Modelling Geomorphological Systems.* ed. M.G. Anderson, pp. 421-440. New York: John Willey & Sons.

Band, L.E. 1986. "Topographic Partition of Watersheds with Digital Elevation Models." *Water Resources Research*, vol. 22, no. 1, pp. 15-24.

Barsalou, L.W. 1983. "*Ad-hoc* Categories." *Memory and Cognition*, vol. 11, no. 3, pp. 211-227.

Carter, J.R. 1989. "Relative Errors Identified in USGS Gridded DEMs." *Proceedings of Auto-Carto 9*, pp. 255-265.

Chorley, R.J., and P. Haggett. 1967. *Physical and Information Models in Geography.* London: Methuen.

Chou, Y.H. 1992. "Slope-Line Detection in a Vector-Based GIS." *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 2, pp. 227-233.

Clarke, K.C. 1988. "Scale-Based Simulation of Topographic Relief." *The American Cartographer*, vol. 15, no. 2, pp. 173-181.

Clarke, K.C. 1990. *Analytical and Computer Cartography.* Englewood Cliffs, New Jersey: Prentice-Hall.

Clarke, K.C., and D.M. Schweizer. 1991. Measuring the Fractal Dimension of Natural Surfaces Using a Robust Fractal Estimator. *Cartography and geography Information Systems*, vol. 18, no. 1, pp. 37-47.

Cromley, R.G. 1992. *Digital Cartography.* Englewood Cliffs, New Jersey: Prentice-Hall.

De Floriani, L., B. Falcidieno, G. Nagy, and C. Pienovi. 1984. "A Hierarchical Structure for Surface Approximation." *Computer & Graphics*, vol. 8, no. 2, pp. 183-193.

Douglas, D.H. 1986. "Experiments to Locate Ridges and Channels to Create a New Type of Digital Elevation Model." *Cartographica*, vol. 23, no. 4, pp. 29-61.

Douglas, D.H., and T.K. Peucker. 1973. "Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or Its Caricature." *The Canadian Cartographer*, vol. 10, no. 2, pp. 112-122.

Ehlschlaeger, C. 1989. "Using the A^t Search Algorithm to Develop Hydrologic Models from Digital Elevation Data." *Proceedings of International Geographic Information System (IGIS) Symposium '89*, pp. 275-281.

Erickson, W.K., and W.C. Likens. 1984. "An Application of Expert Systems Technology to Remotely Sensed Image Analysis." *IEEE 1984 PECORA IX Symposium*, pp. 258-276.

Evans, I.S. 1972. "General Geomorphology, Derivations of Altitude and Descriptive Statistics." *Spatial Analysis in Geomorphology*, ed. R.J. Chorley, pp. 17-90. London: Metheun.

Fairfield, J., and F. Leymarie. 1991. "Drainage Networks from Grid Digital Elevation Models". *Water Resources Research*, vol. 27, no. 5, pp. 709-717.

Firebaugh, M.W. 1988. *Artificial Intelligence.* Boston: Boyd & Fraser Publishing Company.

Frank, A., B. Palmer, and V. Robinson. 1986. "Formal Methods for the Accurate Definition of Some Fundamental Terms in Physical Geography." *Proceedings, Second International Symposium on Spatial Data Handling*, pp. 583-599.

Gerrard, A.J. 1981. *Soils and Landforms.* London: Geogra Allen & Unwin.

Goodchild, M.F., and D.M. Mark. 1987. "The Fractal Nature of Geographic Phenomena." *Annals of the Association of American Geographers*, vol. 77, no. 2, pp. 265-278.

Goodenough, D.G., M. Goldberg, G. Plunkett, and J. Zelek. 1987. "An Expert System for Remote Sensing." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 25, no. 3, pp. 349-359.

Gove, P.B., editor in chief. 1976. *Webster's Third New International Dictionary of the English Language, Unabridged.* Springfield, Maryland: G & C Merriam Co.

Haggett, P. 1975. *Geography: A Modern Synthesis.* New York: Harper & Row, Publishers.

Haralick, R.M. 1983. "Ridge and Valley on Digital Images." *Computer Graphics, and Image Processing*, vol. 22, no. 3, pp. 169-178.

Harvey, P.D.A. 1980. *The History of Topographical Maps*, London: Thames and Hudson Ltd.

Hobson, R.D. 1972. "Surface Roughness in Topography: A Quantitative Approach." *Spatial Analysis in Geomorphology*, ed. R.J. Chorley, pp. 221-246. London: Methuen.

Jenson, S.K. 1985. Automatic Derivation of Hydrologic Basin Characteristics from Digital Elevation Model Data. *Auto-Carto 7 Proceedings*, pp. 301-310.

Jenson, S.K., and J.O. Domingue. 1988. "Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis." *Photogrammetric Engineering and Remote Sensing*, vol. 54, no. 11, pp. 1593-1600.

Johnsonbaugh, R. 1984. *Discrete Mathematics*, New York: Macmillan Publishing Company.

Kulmer, M.P. 1992. *An Intensive Comparison of TINs and DEMs*, Ph.D. dissertation, Department of Geography, University of California at Santa Barbara.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things*, Chicago: The University of Chicago Press.

Labov, W. 1973. "The Boundaries of Words and Their Meanings." *New Ways of Analyzing Variations in English*, ed. J. Freeman, pp. 340-373. Washington, D.C.: Georgetown University Press.

Lam, N.S.N. 1983. "Spatial Interpolation Methods: a Review." *American Cartographer*, vol. 10, no. 4, pp. 308-323.

Lam, N.S.N., and L. De Cola. 1993. *Fractcal in Geography.* Englewood Cliffs, New Jersey: Prentice-Hall.

Lammers, R.B., and L.E. Band. 1990. "Automating Object Representation of Drainage Basins." *Computers & Geoscience*, vol. 16, no. 6, pp. 787-810.

Lay, J.G. 1987. *A Study on Generating Digital Elevation Models (DEM) from Contour Data.* Unpublished M.A. thesis, Department of Geography, University of Hawaii at Manoa.

Lee, D.T., and B.J. Schachter. 1980. "Two Algorithms for Constructing a Delaunay Triangulation." *International Journal of Computer and Information Sciences*, vol. 9, pp. 219-242.

Lee, J. 1989. "A Drop Heuristic Conversion Method for Extracting Irregular Network for Digital Elevation Models." *GIS/LIS '89 Proceedings*, pp. 30-39.

Lee, J. 1991 "Comparison of Existing Methods for Building Triangulated Irregular Network Models of Terrain from Grid Digital Elevation Models". *International Journal of Geographical Information Systems*, vol. 5, no. 3, pp. 267-285.

Lee, J., and P.K. Snyder. 1991. "Modelling Spatial Patterns of Digital Elevation Errors for Drainage Network Analysis." *GIS/LIS Proceedings '91*, pp. 71-79.

Legates, D.R., and C.J. Willmott. 1986. "Interpolation of Point Values from Isoline maps." *American Cartographer*, vol. 13, no. 4, pp. 308-323.

Mandelbrot, B.B. 1967. "How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension." *Science* vol. 156, pp. 636-638.

Mandelbrot, B.B. 1983. *The Fractal Geometry of Nature.* New York: W.H. Freeman and Company.

Mark, D.M. 1975. "Geomorphometric Parameters: A Review and Evaluation." *Geografiska Annaler*, vol. 57(A), no. 3-4, pp. 165-177.

Mark, D.M. 1983. "Relations between Field-Surveyed Channel Networks and Map-Based Geomorphometric Measures, Inez, Kentucky." *Annals of the Association of American Geographers*, vol. 73, no. 2, pp. 358-372.

Mark, D.M. 1984. "Automated Detection of Drainage Networks from Digital Elevation Models." *Cartographica (Auto-Carto Six Selected Papers)*, vol.21, no. 2&3, pp. 168-178.

Marks, D., J. Dozier, and J. Frew. 1984. "Automated Basin Delineation form Digital Elevation Data." *Geo-Processing*, vol. 2, no. 3, pp. 131-141.

McKeown, D.M. 1987. "The Role of Artificial Intelligence in the Integration of Remotely Sensed Data with Geographic Information Systems." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 25, no. 3, pp. 330-348.

Morrison, J.L. 1980. "Computer Technology and Cartography Change." *The Computer in Contemporary Cartography*, ed. D.R.F. Tayler, pp. 5-23. New York: Wiley.

Monkhouse, F.J. (1970) *A Dictionary of Geography.* London: Edward Arnold Ltd.

O'Callaghan, J.F., and D.M. Mark. 1984. "The Extraction of Drainage Networks from Digital Elevation Data." *Computer Vision, Graphics, and Image Processing*, no. 28, pp. 323-344.

Palmer, B. 1984. "Symbolic Feature Analysis and Expert System." *Proceeding of the International Symposium on Spatial Data Handling*, pp. 467-478.

Peucker, T.K., and D. Douglas. 1975. "Detection of Surface-Specific Points by Parallel Processing of Discrete Terrain Elevation Data." *Computer Graphics and Image Processing*, vol. 4, no. 4, pp. 375-387.

Peucker, T.K., R.J. Fowler, J.J. Little, and D.M. Mark. 1978. "The Triangulated Irregular Network." *Proceedings, Digital Terrain Models (DTM) Symposium.* pp. 516-540.

Riazanoff, S., B. Cervelle, and J. Chorowicz. 1988. "Ridge and Valley Line Extraction from Digital Terrain Models." *International Journal of Remote Sensing*, vol. 9, no. 6, pp. 1175-1183.

Robinson, V.B., A.U. Frank, and M.A. Blaze. 1986. " Expert System and Geographic Information Systems: Review and Prospects." *Journal Surveying Engineering*, vol. 112, no. 2, pp. 109 -118.

Rosch, E. 1975. "Cognitive References Points." *Cognitive Psychology*, vol. 7, pp. 532-547.

Rosch, E. 1978. "Principles of Categorization." *Cognition and Categorization*, eds. E. Rosch and B. Lloyd, pp. 28-48. Hillsdale, New York: John Willey & Sons.

Smith, T.R., C. Zhan, and P. Gao. 1990. "A Knowledge-Based, Two-Step Procedure for Extracting Channel Networks from Noisy DEM Data." *Computers & Geosciences*, vol. 16, no. 6, pp. 777-786.

Sprunt, B. 1972. "Digital Simulation of Drainage Basin Development." *Spatial Analysis in Geomorphology*, ed. R.J. Chorley, pp. 371-389. London: Metheun.

Stamp, D. ed. (1966) *Longmans Dictionary of Geography*, ed. London: Longmans, Green & Co Ltd.

Theobald, D.M. 1989. "Accuracy and Bias Issues in Surface Representation." *Accuracy of Spatial Databases*, eds. M. Goodchild and S. Gopal, pp. 99-106. New York: Taylor and Francis.

Tsai, V., and A.P. Vonderohe. 1991. "A Generalized Algorithm for the Construction of Delaunay Triangulations in Euclidean N-Space." *GIS/LIS '91 Proceedings*, pp. 562-571.

U.S. Army CERL. 1992. *GRASS Reference Manual, version 4.0.*

USGS. 1983a. *USGS Digital Cartographic Data Standards. Digital Elevation Models.* Geological Survey Circular 895-B.

USGS. 1983b. *USGS Digital Cartographic Data Standards. Digital Line Graphs from 1:24000-Scale Maps.* Geological Survey Circular 895-C.

USGS. 1983c. *USGS Digital Cartographic Data Standards. Geographic Names Information System.* Geological Survey Circular 895-F.

Vasiliev, I, S. Freundschuh, D.M. Mark, G.D. Theisen, and J. McAvoy. 1990. "What Is a Map?" *Cartographic, vol. 27, pp. 119-123.*

Warntz, W. 1966. "The Topology of a Socio-Economic Terrain and Spatial Flows." *Paper of the Regional Science Association.* vol. 17, pp. 47-61.

Way, D.S. 1978. *Terrain Analysis.* New York: McGraw-Hill.

Winston, P.H. 1992. *Artificial Intelligence.* Menlo Park, California: Addison-Wesley Publishing Company.

Zadeh, L. 1965. "Fuzzy Sets." *Information and Control*, vol. 8, pp. 338-353.