CENTRE FOR ADVANCED SPATIAL ANALYSIS
Working Paper Series

CASA

Paper 43

SURFACE NETWORKS

Sanjay Rana
Jeremy Morley

UCL

Centre for Advanced Spatial Analysis
University College London
1-19 Torrington Place
Gower Street
London WC1E 6BT

[t] +44 (0) 20 7679 1782
[f] +44 (0) 20 7813 2843
[e] casa@ucl.ac.uk
[w] www.casa.ucl.ac.uk

http//www.casa.ucl.ac.uk/working_papers/paper43.pdf

Date: February 2002

ISSN: 1467-1298

Abstract

The desire to understand and exploit the structure of continuous surfaces is common to researchers in a range of disciplines. Few examples of the varied surfaces forming an integral part of modern subjects include terrain, population density, surface atmospheric pressure, physico-chemical surfaces, computer graphics, and metrological surfaces.

The focus of the work here is a group of data structures called Surface Networks, which abstract 2-dimensional surfaces by storing only the most important (also called fundamental, critical or surface-specific) points and lines in the surfaces. Surface networks are intelligent and "natural" data structures because they store a surface as a framework of "surface" elements unlike the DEM or TIN data structures. This report presents an overview of the previous works and the ideas being developed by the authors of this report. The research on surface networks has four main focus areas namely, data structure model, automated extraction, generalisation, and applications. The report is also organised into these research themes.

Despite their immense analytical potential, there have been a number of limitations to date, which need to be tackled:

- Due to their design requirements, current implementations of Surface networks have been restricted to surfaces with fluvial features (i.e., must have ridges, channels, peaks, passes, and pits). However, a number of surfaces have biased topography such as in glaciated or karstic terrains or features may be absent e.g., flat surfaces.

- The feature detection methods are scale dependent. In other words, in any one run, our computing routines detect features that fit into the fixed search window (kernel etc.). An incorrect feature detection method causes loss of the topological

properties, essential for the construction of a consistent surface network.

- Although the topological generalisation of surface networks is well understood, there has been no proposal on the regeneration of the topographical details in the generalised area of the surface networks.

- Surface networks are "believed" to be useful for the visualisation of complex surfaces, optimising visibility and accessibility routines and performing landscape evolution. However, like any other abstraction of surfaces, surface networks also carry a level of uncertainty.

This report describes the results of the research carried out by the reports' authors on the following issues:

- Surface network model: A comprehensive review of the surface network model was done, which revealed some acute limitations of the surface network data model. It was observed that the surface network data model requires significant development to take into account the varied surface forms and the scale issues of terrain data structures.

- Automated extraction: A survey of the algorithms for the automated extraction of surface network revealed that none of the automated extraction methods could extract both a topologically-consistent and complete (taking into account scale-issues) surface network.

- Generalisation: The study of the research on the generalisation of surface networks revealed that the potential of the generalisation is hardly addressed. This work has proposed some alternative methods for the generalisation of surface networks.

- Applications: A survey of the applications of surface network data structure revealed its use in the computer science field mainly for visualisation. This work proposes the use of surface network for optimising viewshed computation and surface evolution studies.

A platform has been to set up to conduct experiments and further investigation on surface networks.

Table of Contents

Acknowledgements.. **iii**

List of Figures.. **iv**

List of Tables.. **vii**

Chapter 1 Introduction **1 – 9**

1.1 Surface Information Encapsulation.. **1**

1.2 Fundamental issues in Surface Topology.. **4**

1.3 Outline of the Report.. **7**

1.4 Domain of the Report..

Chapter 2 The Design of Surface Networks **10 – 27**

2.1 Origins.. **10**

2.2 Surface Network Data Structures.. **13**

2.2.1 Contour and Slope Lines.. **13**

2.2.2 Hills and Dales.. **14**

2.2.3 Critical Point Theory.. **17**

2.2.4 Reeb Graph and Contour Trees.. **19**

2.2.5 Surface Network or Pfaltz's Graph.. **20**

2.2.6 Critical Point Configuration Graph.. **24**

2.3 Summary.. **26**

Chapter 3 Extraction of Surface Networks **28 - 40**

3.1 Introduction.. **28**

3.2 Surface Network Extraction Methods.. **29**

3.2.1 Manual Extraction.. **30**

3.2.2 Triangulation.. **31**

3.2.2.1 Definitions and Methodology.. **31**

3.2.2.2 Discussion.. **34**

3.2.3 Polynomial Surface Fitting.. **35**

3.2.3.1 Definitions and Methodology.. **35**

3.2.3.2 Discussion.. **39**

3.3 Next Research Aims.. **40**

Chapter 4 Extraction of Surface Networks **41 - 59**

4.1 Generalisation.. **41**
4.2 Homomorphic Contraction of Surface Networks.................... **44**
4.2.1 Criteria for Homomorphic Contraction................................ **46**
4.3 Non-Homomorphic Contraction of Surface Networks... **51**
4.4 Generalisation Experiment.. **52**
4.4.1 Methodology.. **52**
4.4.2 Surface Topology Toolkit.. **54**
4.4.3 Results.. **55**
4.5 Regeneration of Surface... **56**
4.6 Discussion.. **59**

Chapter 5 The Applications and Conclusions **60 - 67**

5.1 Scope of Applications.. **60**
5.1.1 Enhanced and Intuitive Visualisation.............................. **61**
5.1.2 Increased Efficiency.. **62**
5.1.3 Simple Surface Evolution.. **64**
5.2 Discussion and Conclusions.. **64**

Bibliography **68**

Acknowledgements

Sanjay would like to thank ACU, UCL and CASA for the funding on this report. Surface network is an interdisciplinary subject and we sought advice from researchers from a variety of subjects with useful feedbacks and materials. The researchers are Chandrajit Bajaj, David Mark, David O' Sullivan, Gert Wolf, Gunho Sohn, Jan Koenderink, Jason Dykes, Joes Staal, John Pfaltz, Jon Cameron, Jo Wood, Muki Haklay, Paul Rosin, Paul Scott, Shigeo Takahashi, Thomas Poiker, and Yosihisha Shinagawa. We thank them al very much.

List of Figures

Figure 1.1	Different representations of the topography around Hoover Dam. (a) Raster or Grid, (b) Contour, (c) Triangulated Irregular Network (TIN), and (d) Surface Network. In the surface network, red dots are peaks, green dots are passes, black and white dots are pits, blue lines are channels, and yellow lines are ridges Terrains have been hill-shaded and coloured by elevation...	**3**
Figure 2.1	(a) A perspective view and (c) contour map of an island with its critical points, and its (b) Reeb Graph, (c) surface network and (d) Pfaltz's Graph (The numbrs indicate the weights. Note than (a) also shows the reduction of contours into the peaks (summits) and pits (minima) as explained by Cayley (1859) and Maxwell (1870).............................	**21**
Figure 2.2	Violation of rule P5a and P5b..	**22**
Figure 2.3	(a) Channel junction and (b) Ridge bifurcation.	**24**
Figure 2.4	Two examples of slope districts – (a) and (c) and their Critical Point Configuration Graph – (b) and (d) respectively. (e) and (f) are the other two basic types of CPCG...	**25**
Figure 3.1	Point $p(i,j)$ in a grid (data view) and its 8 surrounding neighbours..	**31**
Figure 3.2	Point p in a grid (analytical view) and its 7 adjacent neighbours (hollow circles)...	**32**

Figure 3.3 Decomposition of a degenerate pass (Modified from Takahashi et. al, 1995). Figure shows the neighbours and their heights. Higher neighbours are placed inside a grey region. (a) The original neighbour list, (b) the reduced neighbour list, (c) the list in the first turn of the loop in the algorithm, and (d) the final set of neighbours which will define the pass………………………………………………………………. **33**

Figure 3.4 Critical points of the surface and the configuration of their eigenvalues and eigenvectors. R_1 and R_2 are the real parts of the eigenvalues………………………. **37**

Figure 3.5 (a) Elliptic, (b) hyperbolic and (c) parabolic conic sections with their semi-axes idetified (After Wood, 1998). A planar case is not considered here………….. **39**

Figure 3.6 Three possible intersection cases between (circular) region of interest and conic section's semi-axes (After Wood, 1998). (a) Two axes intersect with region – pits, peak or pass, (b) one axis intersects with region – channel or ridge and (c) no intersection – planar……………………………………. **39**

Figure 4.1 Generalisation Lineages and Stages in spatial data structures………………………………………………………………. **42**

Figure 4.2 A hypothetical island and the (y_o – z_o)-contraction and (y_o – z_o)-contraction of its surface network. The numbers in the square brackets in (a) denote the height of the critical points and x is the surrounding pit………………………………………………………. **46**

Figure 4.3 (a) Hypothetical topographic surface and its (b) surface network. Blue contour represents the surrounding pit………………………………………………………. **53**

Figure 4.4 (a) Topography around the Latschur Mountains in the Carinthia Region, Austria and (b) its surface network………………………………………………………………. **53**

Figure 4.5	Graphical User Interface of the Surface Topology Toolkit application with the controls for the contractions………………………………………………………………	**54**
Figure 4.6	Comparison of the effectiveness for selection of points in the surface network (a) between maximum of elevation difference criterion (b) and maximum of edge length criterion (c). Note that criterion (b) selects a long ridge due to its low drop in elevation (350)…………………………………………………	**57**
Figure 4.7	Comparison of the effectiveness for selection of points in a (a) surface network, between (b) sum of elevation difference criterion and (c) valency criterion, showing how criterion (b) can mislead about the ridge/channel crossings. Numbers at peaks in (a) are sum of elevation differences and their valency (in parentheses)…………………………………	**57**
Figure 4.8	Generation of an artificial valley inside the dotted region of the Latschur surface network……………………	**58**
Figure 5.1	Comparison of the dumbness and uncertainty in spatial data structures………………………………………………	**61**
Figure 5.2	Comparison between the current and surface network enhanced visualisation of the dynamism in the geopotential height over Europe………………………	**63**
Figure 5.3	Intervisibility calculations in a part of Isle of Man, based on (a) each grid point (16335 points) and (b) critical points and lines (3975 points). Lighter to darker colour indicates increasing number of observers. The blue line is the surface network………	**65**

List of Tables

Table 1.1	Sequence of the "interdisciplinary" research on surface networks…………………………………………………………	**8**
Table 3.1	Criteria for classification of critical points in the eight-neighbourhood method…………………………………..	**31**
Table 3.2	Criteria for classification of non-degenerate critical points based on delaunay triangulation…………………..	**32**
Table 3.3	Morphometric Features descibed by second derivatives (After Wood, 1996)………………………………	**36**

Chapter 1

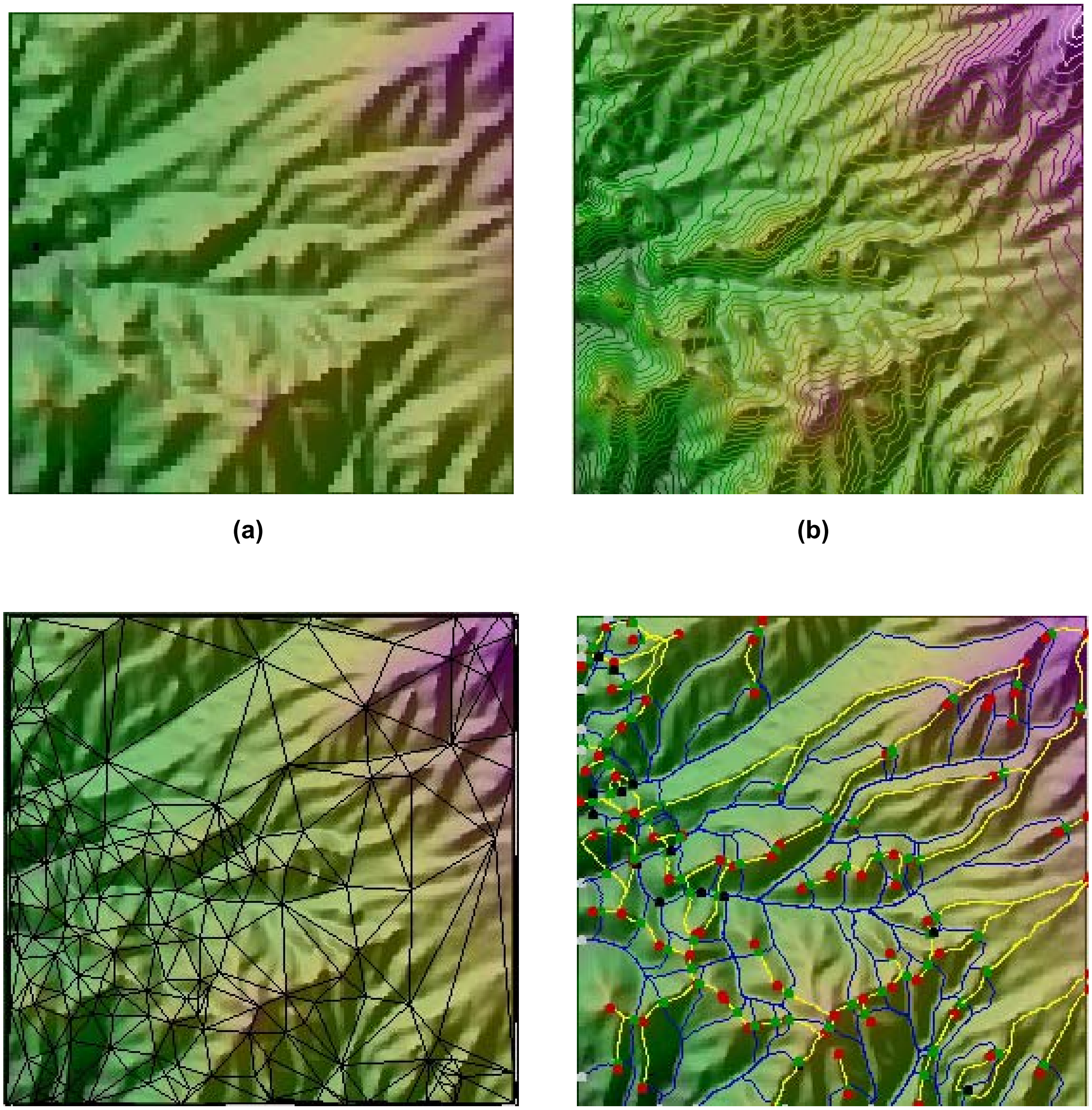
Introduction

1.1 Surface Information Encapsulation

The desire to understand and exploit the structure of continuous surfaces is a common aim to researchers in a range of disciplines. A few examples of the varied surfaces forming an integral part of modern subjects include terrain, population density, surface atmospheric pressure, physico-chemical surfaces, computer graphics, and metrological surfaces. However, with an increasingly multispectral and highly dense data (surfaces in this case), researchers want to be able to filter out redundant observations. These aims (more information but less data volume) seem to contradict each other. However, a right balance between the volume of the data and the information content in the data is an essential requirement to make our analyses (human or robotic) fast and to keep our data storage usage to a minimum. In addition to the understanding, considerable efforts are also spent to produce computing methods to perform data processing automatically.

In general, for practical reasons, more than one kind of data representation is often applied to arrive at a suitable *Information in Data* ($\boldsymbol{D}_{\mathrm{I}}$) to *Data volume* ($\boldsymbol{D}_{\mathrm{V}}$) ratio. Sometimes even layers with different $\boldsymbol{D}_{\mathrm{I}}/\boldsymbol{D}_{\mathrm{V}}$ ratios (i.e., different data structures) are used for the same information. A typical example is the difference between the internal representation of digital images compared to their own optimised file formats. A proper treatment of this issue is beyond the scope of this work but for interested readers literature on data compression, information theory and data structures contains more information. What is important to note here is that there is an inevitable demand for data structure designs and computing algorithms to achieve a satisfactory D_I/D_V.

It will not be an exaggeration to assume that there could be many ways to achieve a suitable D_I/D_V for surfaces. The focus of this work is a data structure, which achieves a D_I/D_V encapsulation in surfaces by storing only the most important (also called fundamental, critical, surface-specific) points and lines in the surfaces. In the Geographic Information (GI) science, such important points and lines have been variously named as landform elements (Speight, 1976), surface specific features (Fowler and Little, 1979), symbolic surface features (Palmer, 1984), surface patches (Feuchtwanger and Peucker, 1987), critical surface features (Wolf, 1992), and specific geomorphological elements (Tang, 1992) amongst others. Most other subjects use the words "critical points" and "critical lines". This work will use the words "critical points" and "critical lines" to represent these features. Like the names, there have been many proposals for the list of the most important points and lines. However, the peaks (local maxima), passes (local saddles), and pits (local minima) are considered to be the simplest and sufficient set of points to characterise the surface. The topological framework of the surface is constructed with the addition of critical lines, which connect the critical points. The critical lines are ridges (lines linked from peaks to passes), and channels (lines linked from passes to pits). These kinds of data structures are used extensively in various disciplines with different names and construction. Some prominent types of these data structures include the Surface Network, Surface Tree, Critical Point Configuration Graph, Reeb Graph and the many unnamed ones. Fig. 1.1 gives an example of the different representation of the topography in an area around the Hoover Dam, USA.



(c) (d)

Figure 1.1 Different representations of the topography around Hoover Dam. (a) Raster or Grid, (b) Contour, (c) Triangulated Irregular Network (TIN), and (d) Surface Network. In the surface network, red dots are peaks, green dots are passes, black and white dots are pits, blue lines are channels and yellow lines are ridges. Terrains have been hill-shaded and coloured by elevation.

As mentioned earlier, these varieties of the critical point-critical line data structures could have different construction but their “Surface Topology” has the same set of components. Thus, based on this similarity between these data structures and for the sake of simplicity, we propose here author the following terms are used,

- “Surface Network” for the spatial representation, and
- “Surface Network Graph” for the graph representation,

for all data structures constructed with critical points and critical lines. Note that this definition excludes Triangulated Irregular Networks (TIN) because they contain both the critical and ordinary set of points in their structure. The above-mentioned convention will be used in the following parts of this report. However, the author realises that this proposal can only be sensible if there were to be a universal standard on the structure and implementation of surface networks. A specific aim of this research is to combine the aims and methods of various disciplines on this subject.

1.2 Fundamental issues in Surface Topology

There are many other types of surface topological data structures in GI science and other subjects. Wolf (1993) has given a review of some prominent surface topological data structures. In order to achieve a thorough grounding for this work, it is essential to define rigorously those aspects of surface topology and topological data structures which are used to describe surfaces.

Q1. Why should we have surface topology based data structures?

- The data that will be needed to define the surface will be very much reduced. The reduction in size could be as much as 90% (Helman and Hesselink, 1991).

- Topological connections are a much more efficient way to access a spatial database. In this case, surface networks provide a more natural and thus intuitive control on the structure of the surfaces.
- Components in a topological data structure are interdependent and linked. Thus, these data structures can be used for applications that require uniform and controlled response from the entire surface such as morphing in computer graphics and erosion modelling. For example, in the case of surface (represented as contours) generalisation, a common problem in approaches based on line-simplification is the intersection of contours after simplification. However, in the case of a surface network representation of the surface, the use of formal topological simplification prevents the generation of an unrealistic surface after generalisation (Wolf, 1984).
- Surface topology is found more useful for the visualisation of surfaces especially 3D surfaces. It is because as it does not involve the complications of deciding the appropriate colour mapping or contour interval or the density of triangles (in case of TIN) (Helman and Hesselink, 1991; Bajaj and Schikore, 1996).
- As the surface networks are translation- and rotation- invariant they also forms an ideal mechanism for correlating and co-registering surfaces (Bajaj and Schikore, 1996).

Q2. What should the data structures for the surface topology attempt to describe?

Wood (1996) posed a more general form of this question about the extent of characterisation possible for landscape. Wood (1996) remarked that an objective identification of the 'true' landscape is not possible without a stricter definition of terms. It is because of the reason that the concept of 'landscape' is subjective not only to the physical geomorphological process, but is also defined by its use and the preconceptions of the observer. Since surface topology is also a characterisation of the surface therefore its shares these limitations.

- A common expectation from the topological data structures is that they should provide a unified global description of surface. A global description would ensure a sympathetic response in the whole surface if a change occurs in one part of the data structure. Thus, giving a formal control on the continuity of materials and processes that exists in nature as well.
- The data structure should be able to represent most surfaces i.e., both fluvial and non-fluvial (with no or incomplete set of pits, peaks, and passes).
- Though not a necessity it should have the flexibility of undergoing topological adjustments with formal routines such as needed for generalisation of terrain.

Q3. Which surface features should be considered as most important to be included in the surface topology data structures?

Various surface specific features have been proposed to represent the surface. The choice was largely based on the specific applications for which the surface was being modelled. The choice of surface specific features for the framework of surface topology is very essential, as it will decide the following important factors:

- Resemblance to a real surface: The combination of surface specific features selected for surface topology should be able to describe most of the surface forms. However, the more detailed set of surface specific features are selected, the more difficult it will be to handle the data structures.
- Potential applications based on surface topology: As stated earlier, it will be desired that the description of surface topology (in terms of the relations of the selected surface specific features) should contain adequate ways by which properties of surfaces such as drainage networks (hydrological applications), and others can be derived.

1.3 Outline of the Report

Surface Networks have received intensive research inputs from researchers especially in computer science (vision, graphics), geographic information science (terrain modellers) and, to a limited extent, by social scientists. The research on surface networks can be broadly divided into four main areas namely the *Design* (i.e., data structure model), *Extraction* (Automated, Digital), *Generalisation*, and *Applications*. The chronological sequence of research in these areas is shown in Table 1.1. The report is also divided into four parts based on the research areas, namely Theoretical or Design of Surface Networks (Chapter 2), Extraction (Chapter 3), Generalisation (Chapter 4), and the Applications (Chapter 5). These chapters are mostly self-contained description on these areas and include a conclusion either during the description or at the end of the chapters. This has been done to ensure a consistency of thoughts.

Chapter 2 presents a review of the various surface network data structures with insights into their design and applications. This chapter will propose properties expected in a general design of surface network in order to be applicable for most kinds of two-dimensional surfaces.

Chapter 3 focuses on the automated extraction of surface networks. This will involve the treatment of issues such as scale, feature identification and generation of a consistent topology. Three techniques of the extraction namely manual, triangulation and surface fitting will be described in details.

Chapter 4 describes the simplification or technically the generalisation of surface networks. It will explain the importance measures (weights), generalisation criteria and will show the results of the generalisation experiments on a real and a hypothetical terrain. Based on empirical observations, it also proposes new methods for generalisation and

Researcher(s)	Surface Network Research Areas: Design	Extraction	Generalisation	Visualisation
Reech (1858)*	x	--	--	--
Cayley (1859)^	x	--	--	--
Maxwell (1870)^	x	--	--	--
Morse (1925)*	x	--	--	--
Reeb (1946)*	x	--	--	--
Warntz (1966)^	x	--	--	--
Morse (1966)^	x	--	--	--
Pflatz (1976)*^	x	x	x	x
Mark (1977)^	x	?	x	--
Nackman (1984)*	x	x	--	--
Wolf (1984)^	x	--	x	x
Takahashi et. al (1995)*^	x	x	--	x
Rana (2000)^	--	--	x	x
Biasotti et. al (2000)*^	x	x	--	x
Wood, Rana (2000)^	x	x	x	x

Table 1.1 Sequence of the "interdisciplinary" research on surface networks. **^** indicates a research based on terrains as the example of surface and * indicates a research based on a mainly mathematical treatment of surfaces.

comments on the regeneration of the surface around the topological adjustments.

Chapter 5 is largely a demonstration of the ideas and techniques developed in the previous sections. Case studies on the use of surface networks for terrains and meteorological surfaces for visibility analysis and visualisation are described. This chapter concludes with a brief summary of the report and presents the directions for future research.

1.4 Domain of the Report

Until now in the report, the term surface was used loosely to indicate continuous surfaces in n-dimensions. However, in this research a surface has the following strict definition:

- Surface is a twice continuously differentiable function, whose each point (x,y) is associated with its scalar property i.e., $z = f(x,y)$ and
- It is defined over a domain, which is simply connected and bounded by a closed contour line, therefore there are no holes in the surface.

Chapter 2
The Design of Surface Networks

2.1 Origins

From early in various academic fields, attempts have been made to parameterise surfaces into frameworks woven around the geometrical and topological relationships of the fundamental features of the surface. Efforts have taken place in disciplines such as physical and social geography, computer science (particularly graphics and vision), medical sciences, metrology, physics and others, in which the data and output is often a continuous surface. The aim of this chapter is to discuss the various surface network data structures especially around their capabilities to represent the surfaces accurately. In the following text a brief description on the sequence of events related to the developments of surface networks will be given. The details on the individual events are provided thereafter.

The most crucial thought, which was instrumental in the surface network field, was the recognition of the fundamental features. Fundamental features are characteristic features, which are common to all surfaces and contain sufficient information to construct the whole surface, thus taking away the need to store each point on the surface.

Mark (1977) reported that Reech (1858) was perhaps the first to discuss the critical points on a closed surface. It was soon followed by Cayley (1859), who proposed the subdivision of topographic surface into a framework of summits, immits, knots, ridge lines and course lines.

Maxwell (1870), based on purely empirical observations about terrains, proposed relations between the number of summits, number of passes, number of immits (also called bottoms) and number of bars. He also described the partition the topographic surface into "Hills and Dales" based on these features.

In contrast, Morse (1925) proved the same relations between the number of peaks (summits), number of passes (bars), and number of pits (immits) based on differential topology. In general, Morse proposed formal relations between the critical points in an n-dimensional surface, which is known as the Critical Point Theory or Morse Theory. The generic nature and wide applicability of Morse Theory led to the expansion in the interest in the critical points of surfaces amongst various disciplines.

In a significant related development, Reeb (1946) proposed representing the splitting and merging of equi-height contours (i.e., a cross-section) of a surface as a graph. Now as the contours close at the pits and the peaks, and split at the passes, therefore the vertices of this graph, now called Reeb Graph, are the critical points of the surface. The edges of the Reeb Graph turn out to be the ridges and channels. The Reeb Graph was particularly useful because unlike the description of the relationships critical points on the surface given in the Morse Theory, it addressed the embedding of the critical points on the surface.

Warntz (1966) revived the interest of geographers and social science researchers into critical points and lines when he applied the "Hills and Dales" idea for socio-economic surfaces, referred to as the Warntz Network (Mark, 1977).

Another interesting representation of topological relationships between the critical points of a surface is the Contour Tree (Morse, 1968, 1969). Contour Tree represents the adjacency relations of contour loops. The tree like hierarchical structure develops due to the fact that each contour loop can enclose many other contour loops but it can itself be enclosed by only one contour loop. As is evident the Contour Tree is same as the Reeb Graph except separated by two decades. Interestingly, Kweon and Kanade (1994) proposed another similar idea called the Topographic Change Tree. Are these examples of duplicate researches?. It looks so because the bibliography of Kweon and Kanade (1994) does not mention about the Contour Tree work while Mark (1977), who discussed Contour

Tree in details, does not mention Reeb Graphs. As in the case of Reeb Graph, the vertices of such a contour tree are the peaks, pits and passes.

After about a decade Pfaltz (1976) combined Morse Theory inequalities and Warntz Network in a formal graph-theoretic data structure called Surface Network (also called Pfaltz's Graph - coined by Mark, 1977). Since he was in the computer science field, his work attracted the attention of researchers in three-dimensional surfaces such as in medical imaging, crystallography (Johnson et. al, 1999; Shinagawa et. al., 1991) and computer vision (Koenderink and Doorn, 1979). Pfaltz also proposed a graph-theoretic method called homomorphic contraction for generalising the Pfaltz's graph and made the first attempt at the automated generation of surface networks.

Mark (1977) proposed a pruning of the contour tree to remove the nodes (representing contour loops) which do not form the critical points, i.e., the vertices, of the contour tree, and called the resultant structure "Surface Tree". This essentially reduces the contour tree to the purely topological state of a Pfaltz's graph. It is easy to realise that the Reeb Graph, Pfaltz's Graph and Surface Tree have fundamental similarities and are actually inter-convertible (Takahashi et. al, 1995).

Nackman (1984) proposed a new construction for the graphs of critical points, called the "Critical Point Configuration Graph (CPCG)", to be a surface network under more general conditions than those in the Pfaltz's graph. In most simple terms, the CPCG is made up of four basic combinations of the critical points called the Slope Districts (areas of overlap between Hills and Dales).

The next major work in the Pfaltz's Graph lineage was by Wolf (1984), who introduced more topological constraints for the Pfaltz's graph to be a consistent representation of the terrain. He proposed assigning weights to the critical points and lines to indicate their importance in the surface and thus he proposed the name "Weighted Surface Network" for the Pfatlz's graph. He demonstrated new weights-based criteria and methods for the contraction of the surface networks. He, however, performed a manual extraction of surface network from contour maps.

Feuchtwanger and Poiker (1987) proposed a topological model for terrains, which was essentially a combination of ideas from the

Interlocking Ridge and Channel Network (Werner, 1988), Hills and Dales, Contour Tree, Surface Tree, and Pfaltz's Graph. Sadly, although interesting, the idea did not advance beyond the Entity-Relationship Model of the data structure.

A major contribution in surface networks came from Takahashi et. al (1995), who combined the Morse Theory and the Reeb Graph ideas and proposed robust algorithms for the automated extraction of a consistent Surface Network from DEM. A unique aspect of his work was that he used a triangulation based feature detection method to extract the critical points.

On the contrary, Wood (1998), Wood and Rana (2000) attempted to extract the critical points using a polynomial based feature extraction technique with limited success. The advantage of a polynomial-based detection is that it could be adjusted to extract features at various scales unlike the triangulation-based technique, which is restricted to a fixed scale. Rana (2000) discussed the characteristics of Wolf (1984)'s generalisation criteria and proposed an arbitrary user-defined contraction for surface networks.

Now the stage has been set up to describe each of the above-mentioned work in details. The work of Reech (1858) was not available to the author at the time of writing this report so it will not be discussed.

2.2 Surface Network Data Structures

2.2.1 Contour and Slope Lines

Cayley (1859) described the configuration of terrains based on the arrangement of contour lines and slope lines.

Let us assume a mountainous island, the exterior or sea level contour line is therefore a closed curve. There are three main possible configurations of contour lines. A contour line could enclose contour lines of higher elevation or lower elevation or meet contour lines of equal elevation. The contour line bounding an elevation would gradually get smaller and ultimately reduce to a point, which is called a Summit. The contour line bounding a depression would similarly become smaller and reduce to a point, which is called Immit. At some points in the terrain, a contour line may meet three contour lines of the equal elevation. At these points, the surface is horizontal, and one descends in the backward and forward directions while the other ascends in right and left directions. These points are called Knots.

The indicatrix at a summit and immit is an ellipse except in the case when the summit or immit is an umbilicus – the indicatrix then is a circle. Hence in the case of an elliptic indicatrix, all slope lines except one intersect direction of least curvature of the ellipse. The remaining contour lines intersect the contour lines of maximum curvature of the ellipse. The indicatrix at a knot is a hyperbola and therefore the contour lines in the neighbourhood of a knot are similar and similarly situated concentric hyperbolas. At the knot, there are two orthogonal slope lines, which bisect two opposite contour line hyperbolas. This pair of slope lines is the Ridge and Course lines. A knot is a point of minimum elevation for a ridge line while it is a point of maximum elevation for a course line. A ridge line would reach from a knot to summit and a course line would reach from an immit to another immit via a single intervening knot. However, the course line can also arrive at the sea-level contour without reaching another immit. The ridge line or course line may start and end at the same summit or immit respectively, thus forming a closed curve.

2.2.2 Hills and Dales

Maxwell (1870) developed the Cayley (1859) description of the Surface Topology of terrains. Like Cayley (1859) he proposed his ideas based on an elevated surface surrounded by a depression.

The regions of elevation and depression on the surface define the surface in mainly three ways. Firstly, two regions of depression would expand until they meet up at a point, which is called a Bar. It may happen that more than two regions of depression may meet up such as in the case of monkey saddles, which are called degenerate points, but these points are not included in the hypothesis. Secondly, two regions of depression may send out arms, which may meet each other, thus cut off a region of elevation in the middle of the region of depression. The point of meeting, which is called a Pass, cuts off two regions of elevation from one region of depression. Thirdly, the regions of elevation and depression are finally reduced to points, which are called Summits or Tops and Immits or Bottoms respectively.

Given the above ways of generation of the features, Maxwell (1870) derived relations between the number of summits, passes, immits and bars. Every new region of elevation produces a pass. A summit is produced when every new region of elevation is reduced to a point. Therefore, since the whole surface of the earth is a region of depression, the number of summits, S, is one more than the number of passes, P, i.e.,

$$S = P + 1 \quad \text{....(2.1)}$$

Similarly, with every new region of depression, a bar is produced and an immit develops when the region of depression is reduced to a point. Therefore the number of immits, I, is one more than the number of bars, B., i.e.,

$$I = B + 1 \quad \text{....(2.2)}$$

A pass or a bar can be called a single, double, or n-ple according to two, three, or n+1 regions of elevations or depressions meeting at a point.

He added that these rules apply to any function of two variables. The summits are the maxima and the immits are the minima. Therefore, based on the two eqs. 2.1 and 2.2, for this function with number of maxima, p, and number of minima, q, there are

$$p + q - 2 \quad \text{....(2.3)}$$

cases of stationary values, which are neither maxima nor minima. He extended this relation, in an interesting way, to function of three variables, which is beyond the scope of this report.

Geomorphologically and analytically (as expressed by eq. 2.3), the bars and passes are the same features, called saddles or passes. The points of stationary values which are neither maxima nor minima are in fact the saddles, which gives us the following important relation between the number of summits, S, number of immits, I, and number of passes, P,

$$I - P + S = 2 \quad \text{....(2.4)}$$

It will be shown in section 2.2.3 how this relation can be derived from differential topology.

Slope lines are lines that are everywhere at right angles to the contour lines. All slope lines, except two, when ascending generally reach a summit and when descending end at an immit. The exceptional two slope lines reach a pass or a bar. The surface is divided into two types of Districts (areas of surface). These are the Dales or Basins, whose slope lines converge at the same immit and the Hills, whose slope lines originate at the same summit, which are called Hills. Dales and Hills are partitioned by Watersheds and Watercourses respectively. A watershed can be drawn from a pass or a bar by tracing the slope line from the maxima connected to this pass (bar) until it reaches a summit. A watercourse is similarly a slope line starting from the minima connected to a pass (bar) and ending at an immit. Lines of watershed never reach an immit and lines of watercourse never reach a summit.

Based on the deductions above the total number of summits, S, on the whole surface is

$$S = 1 + p_1 + 2p_2 + \ldots + (n\text{-}1)p_{n\text{-}1} \quad \text{....(2.5)}$$

where p_1 is the number of single passes, p_2 is the number of double passes and so on, and n is the maximum number of regions of elevation meeting up at the summits. The total number of immits, I, is

$$I = 1 + b_1 + 2b_2 + \ldots + (n\text{-}1)b_{n\text{-}1} \quad \text{....(2.6)}$$

where b_1 is the number of single passes, b_2 is the number of double passes and so on, and n is the maximum number of regions of depression meeting up at the immits. Therefore the number of watersheds, W, will be

$$\mathrm{W} = 2\,(b_1 + p_1) + 3\,(b_2 + p_2) + \ldots + (n\text{+}1)\,(b_{n\text{-}1} + p_{n\text{-}1}) \quad \text{....(2.7)}$$

where n is the order of pass and bar i.e., single, double and so on. The number of watercourses is similarly defined.

Now according to the Listing's rule for finding the number of faces,

$$P - L + F - R = 0 \quad \text{....(2.8)}$$

where P is the number of points, L is the number of lines, F is the number of faces and R is the total number of regions. Here $R = 2$, viz. the earth and the surrounding spaces, hence

$$F = L - P + 2 \quad \text{....(2.9)}$$

If we assume that L represents the watersheds, thus P equals to the total number of summits, passes, and bars then F is the number of Dales, which is evidently the number of immits. But we could also assume that L

represents watercourses, then P will equal to the total number of immits, passes and bars and F will be the number of Hills or i.e., the summits. Finally, if we assume that L is equal to the total number of lines, and P is equal to the total number of points then F, the total number of natural districts i.e., the hills and dales together, is equal to the total number of watersheds and watercourses or the total number of summits, immits, passes and bars minus 2.

Warntz (1966) reiterated these ideas and proposed their use in understanding socio-economic surfaces and spatial flows. For an example he used population potential surface of USA and demonstrated various applications of Hills and Dales in transport network density, movement of money etc.

2.2.3 Critical Point Theory

The first purely mathematical treatment of surface networks came from Morse (1925). He considered the "critical points" of a sufficiently smooth function f defined over an arbitrary n-dimensional manifold M where f satisfies appropriate conditions on the boundary of the manifold.

Milnor (1963) is a widely referred book for a background reading on Morse Theory but this book is out of print and is not available easily (and was not available to the author either). Three good alternative sources are Pfaltz (1978), Takahashi (1996) and the Encyclopaedia Britannica.

The conditions and definitions on the surface function f are:

- f is sufficiently smooth if $f \in C^2$ i.e., it has continuous 2nd derivatives. Thus, it is possible to calculate the curvature at each point on the function so cases like overhangs and lakes do not exist,
- A point $p \in M$ is a critical point of f if $\delta f(p) = 0$ i.e., the 1st partial derivative of f vanishes at p or f is "locally flat" at p.
- For all points b on the boundary $f(b) > f(i)$ where i is an interior point.
- All critical points of f are non-degenerate i.e., the matrix $H(f)$ of the 2nd derivatives, called the Hessian Matrix, at a point $p(x,y)$ has a nonzero determinant (i.e., singular or regular). The Hessian matrix for p is defined as

$$H(x,y) = \frac{\delta f}{\delta f_x \delta f_y} \quad ....(2.10)$$

- The index of the critical point p of f is the number of negative eigenvalues of the Hessian matrix at p.

  In this work, the dimension of the manifold is 2, therefore the indices of the critical points are from 0 to 2. It turns out that the peak of the function f has the index 2, a pass has the index 1 and a pit has the index 0.
- The function f on M is called a Morse function if it has no degenerate critical point. An example of a degenerate critical point in terrain is the monkey saddle. A point at a monkey saddle although locally flat has a non-singular Hessian matrix.

With these conditions and premises for the function f and its critical points, Morse related the number of critical points of f with the topology of the Manifold. The details of the comparison are not specifically relevant to be provided here. But the following inequalities derived by him for a 2-dimensional sphere (f is assumed to be a part of the sphere) are important to be noted here:

$$P_0 \geq 1 \quad ....(2.11)$$

$$P_0 - P_1 \geq 1 \quad ....(2.12)$$

$$P_0 - P_1 + P_2 = 2 \quad ....(2.13)$$

where P_0, P_1 and P_2 denote the critical points of index 0,1 and 2 respectively. As mentioned, earlier in the discussion, in the case of 2-dimensional function they correspond to pits, passes and peaks respectively. Note the similarity between the sophisticated eq. 2.13 and the simpler eqs. 2.4 and 2.9. These inequalities are a simple example of Critical Point Theory or Morse Theory by Morse (1925).

In later works, Morse demonstrated the use of these relations in the understanding of various surfaces such as in physics, biology, and economics and thus encouraged the wide spread use of the Morse Theory

However, there are two crucial issues, which Morse did not address. Firstly, he did not establish various possible "configurations" of the critical points within the manifold (Pfaltz, 1978). The following sections will describe some ways of representing the configuration of the critical points.

Secondly, it is well known that surfaces, especially terrains do contain abundant degenerate points. Therefore, terrains are not ideally a Morse function by definition. However, the potential advantages of storing the "structure of a surface" in a critical point framework are very attractive and it will be shown in the next chapter that degenerate points can be hypothetically "decomposed" into a non-degenerate point.

2.2.4 Reeb Graph and Contour Trees

A Reeb Graph (Reeb, 1946) is a graph which represents the splitting and merging of equi-height contours (Takahashi et. al, 1995). The original article by Reeb (1946) is in French but formal discussions of his ideas in English are given by Takahashi et. al (1995), Takahashi (1996) and Biasotti et. al (2000).

The following description of the Reeb Graph is largely taken from Takahashi et. al (1995). For a function f representing the height of a terrain, its Reeb Graph is obtained by identifying points p and q if the two points are contained in the same connected component on the cross-section of the surface at the height $f(p) = f(q)$. Thus, a cross-sectional contour is represented as a point of the edge of the Reeb Graph (Fig. 2.1). As explained in section 2.2.1 contours converge or diverge at the critical points, therefore the vertices of Reeb Graphs represent the critical points of f. Fig. 2.1a shows an example of a mountain and its critical points and Fig. 2.1b is its corresponding Reeb Graph. The combination of the Reeb Graph with the Morse Theory could be one formal way of representing the topological configuration of the critical points on the surface as a single data structure. Biasotti et. al (2000) has developed the Reeb Graph to model terrains, and it is referred as the Extended Reeb Graph (ERG). One of the main characteristic of ERG is that it uses the areas around critical points, called critical areas, as that allows a better reconstruction of the surface. Biasotti's work is an example of widely discussed issue in geomorphometry of whether a peak is actually a point or an area.

It is interesting to note that the construction of Reeb Graph is very similar to the Contour Tree (Morse, 1968; 1969), Surface Tree (Mark, 1977), and Topographic Change Tree (Kweon and Kanade, 1994).

2.2.5 Surface Network or Pfaltz's Graph

Pfaltz (1976) was the first researcher who proposed a formal topological data structure for surfaces based on the combination of the Critical Point Theory and the theory of Hills and Dales. He essentially added the missing connectivity between the critical points of a surface (which is a Morse function) in the Critical Point Theory by using the relationships defined between the critical points in the theory of Hills and Dales. He proposed that the relationships between the critical points can be represented by a tripartite (three sets of critical points) directed graph, which he called the Surface Network, also known as Pfaltz's Graph (Mark, 1977). For example for the surface in Fig. 2.1a, its surface network and Pfaltz's Graph are shown in Fig. 2.1c and Fig. 2.1d respectively. However, not all such tripartite graphs can represent a real surface (Pfaltz, 1976; Wolf, 1984). A weighted, directed, tripartite graph $W = (P_0, P_1, P_2; E)$, where P_0, P_1, P_2 are the three vertex sets representing the sets of all pits, passes and peaks, respectively, while E is the set of all edges, is termed a (weighted) surface network (WSN) if

P0: W is planar.

This means that an intersection of edges for instance an intersection of ridges and channels is not allowed. This is natural because except at the critical points, there can only be one type of slope line passing through one point.

P1: The subgraphs $[P_0, P_1]$ and $[P_1, P_2]$ are connected.

This means that channels connect pits and passes, and ridges connect peaks and passes.

P2: $|P_0| - |P_1| + |P_2| = 2$

It states that the number of pits minus the number of pass points sum the number of peaks must always be two (see section 2.2.2 and section 2.2.3 for the proof).

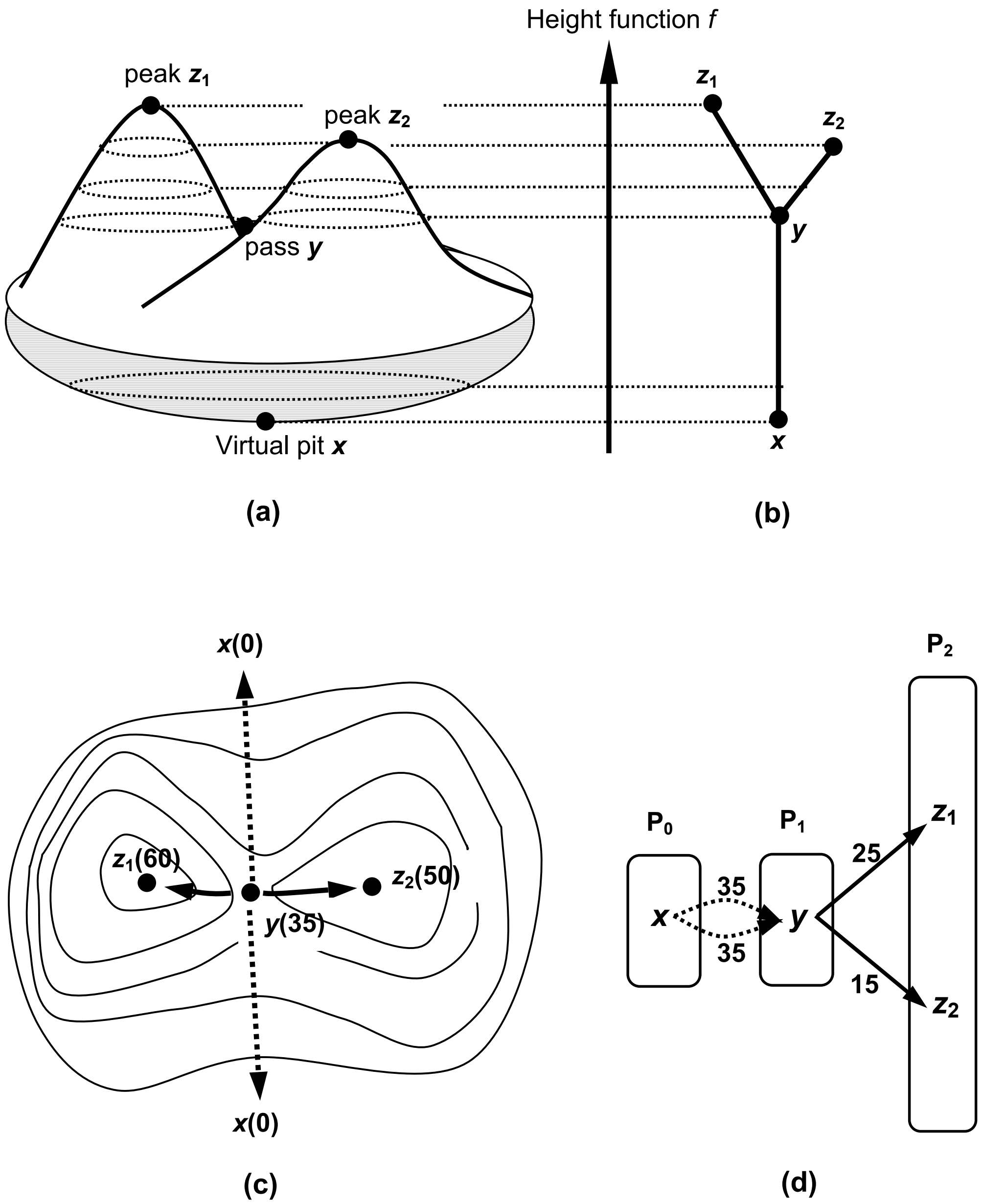


Figure 2.1 (a) A perspective view and (c) contour map of an island with its critical points, and its (b) Reeb Graph, (c) surface network and (d) Pfaltz's Graph (The numbers indicate the weights). Note that (a) also shows the reduction of contours into the peaks (summits) and pits (minima) as explained by Cayley (1859) and Maxwell (1870).

P3: For all $y \in P_1$, $id(y) = od(y) = 2$ where y = pass, $id(y)$ = in-degree of y, $od(y)$ = out-degree of y.
This means that exactly two channels and exactly two ridges emanate thus excluding the existence of degenerate passes. As can be seen in nature, this property is most often violated for example in the case of channel junctions and ridge bifurcations. Pfaltz (1976) suggested that these points could be "decomposed" into normal critical points. Wolf (1990) and Takahashi et. al (1995) proposed solutions, which will be discussed later in this section and in the next Chapter.

P4: $val(x, y_i) = val(y_i, z) = 1$ implies that there exists $y_j \neq y_i$ such that (x,y_j), $(y_j,z) \in E$, where x = pit, y = pass, z = peak and val = valency.
It guarantees that if there is a path from pit x via pass y_i to peak z, which consists only of edges with valency one, then there exists another path from pit x to peak z via a distinct saddle y_j.

P5a: (x,y) is an edge of a circuit in the bipartite graph $[P_0,P_1]$ iff $val(y,z) \neq 2$ for all $z \in P_2$
P5b: (y,z) is an edge of a circuit in the bipartite graph $[P_1,P_2]$ iff $val(x,y) \neq 2$ for all $x \in P_0$
This property asserts that a configuration as shown in Fig. 2.2 is impossible.

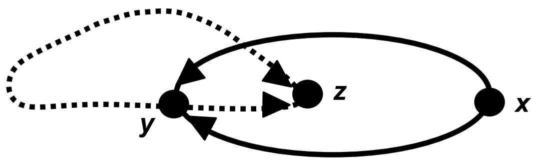


Figure 2.2 Violation of rule P5a and P5b.

P6: $w(e_i) > 0$ for all $e_i \in E$
This means that all the edge weights must be greater than zero. For instance, if $h(x_0)$, $h(y_0)$ and $h(z_0)$ represents the elevations of a pit, pass and peak, respectively, then the weight of a channel is $h(y_0)$ - $h(x_0)$ and the weight of a ridge is $h(z_0)$ - $h(y_0)$.

P7: For all $x \in P_0$, y_i, $y_j \in P_1$, $z \in P_2$ and (x,y_i), (x,y_j), (y_i,z),$(y_j, z) \in E$ holds
$w(x,y_i)+ w(y_i,z) = w(x,y_j)+ w(y_j,z)$
This means that for all paths from pit x to peak z the difference in elevation is the same, no matter which saddle point is passed.

P8a: If $val(x,y) = 2$ with e_{i1} =(x,y) and e_{i2} =(x,y) then $w(e_{i1}) = w(e_{i2})$
P8b: If $val(y,z) = 2$ with e_{i1} =(y,z) and e_{i2} =(y,z) then $w(e_{i1}) = w(e_{i2})$
This means that all channels from a pit to a pass have the same difference in altitude; the same holds for ridges, too.

Wolf (1984) developed Pfaltz's Graph and proposed weights to be assigned to the critical points and lines to indicate their importance in the local or global structure of the surface. He thus called the new form a Weighted Surface Network (WSN). Although surface networks are an abstraction of surfaces, they could still have redundant information. Pfaltz (1976) proposed a graph-theoretic method of simplification of surface networks called Homomorphic Contraction, which can remove redundant vertices and edges but still preserve the above-mentioned topological properties of the surface network. Wolf (1984) developed Pfaltz's ideas on homomorphic contraction and introduced the use of weights and various criteria for the contraction. More information on the contraction is explained in Chapter 4, which describes the generalisation of surface networks.

It is evident from Fig. 2.1c,d that the surface networks are purely a topological data structure. However, as Wolf (1993) commented an ideal data structure for a surface should be able to describe both the topological and geometrical properties of the surface. This issue has been addressed in mainly two ways.

Wolf (1990) proposed the addition of geographic co-ordinates to the critical points, thus the surface network could be triangulated to represent the "topography" of the surface. He termed the new surface network a Metric Surface Network (MSN). With MSN, he was also able to provide solutions for the problem of the absence of representation for the two important topographic points - the channel junctions and the ridge bifurcations. He proposed that channel junction and ridge bifurcation could be represented as an infinitesimally close pair of pit-pass and pass-peak, respectively (Fig. 2.3). In the case of junctions and bifurcations, an arbitrary low weight can be assigned to indicate their proximity, for example Wolf (1990) used a value of 2.

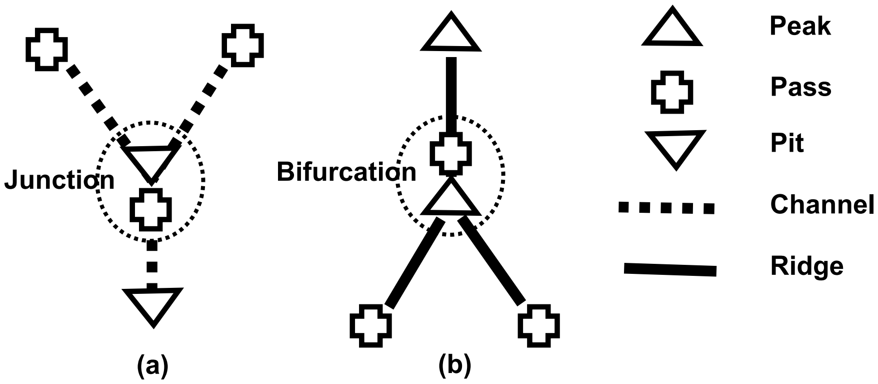


Figure 2.3 (a) Channel junction and (b) Ridge bifurcation.

Takahashi et. al (1995) proposed the use of Reeb Graphs to reconstruct the topography as they store information about the hierarchy of the contours. He found that it was easy to construct the Reeb Graph from the surface network, as it will be very time consuming to detect the topological changes in the cross-sectional contours.

The next chapter will present the slope lines (ridges and channels) based approach (Wood and Rana, 2000) to maintain the topographic appearance and the topological virtues of the surface networks.

2.2.6 Critical Point Configuration Graph

Nackman (1984) also proposed a graph-theoretic based topological data structure, called Critical Point Configuration Graphs (CPCG), for surface (assumed to be a Morse function) based on the combination of the Critical Point Theory and the theory of Hills and Dales. He was motivated by the idea of surface networks (Pfaltz, 1976) but instead of partitioning a surface in a single framework of critical points and lines, he proposed the subdivision of surfaces, especially terrains, into slope districts. Slope districts are regions where Hills and Dales overlap (Fig. 2.4a,c). He proved

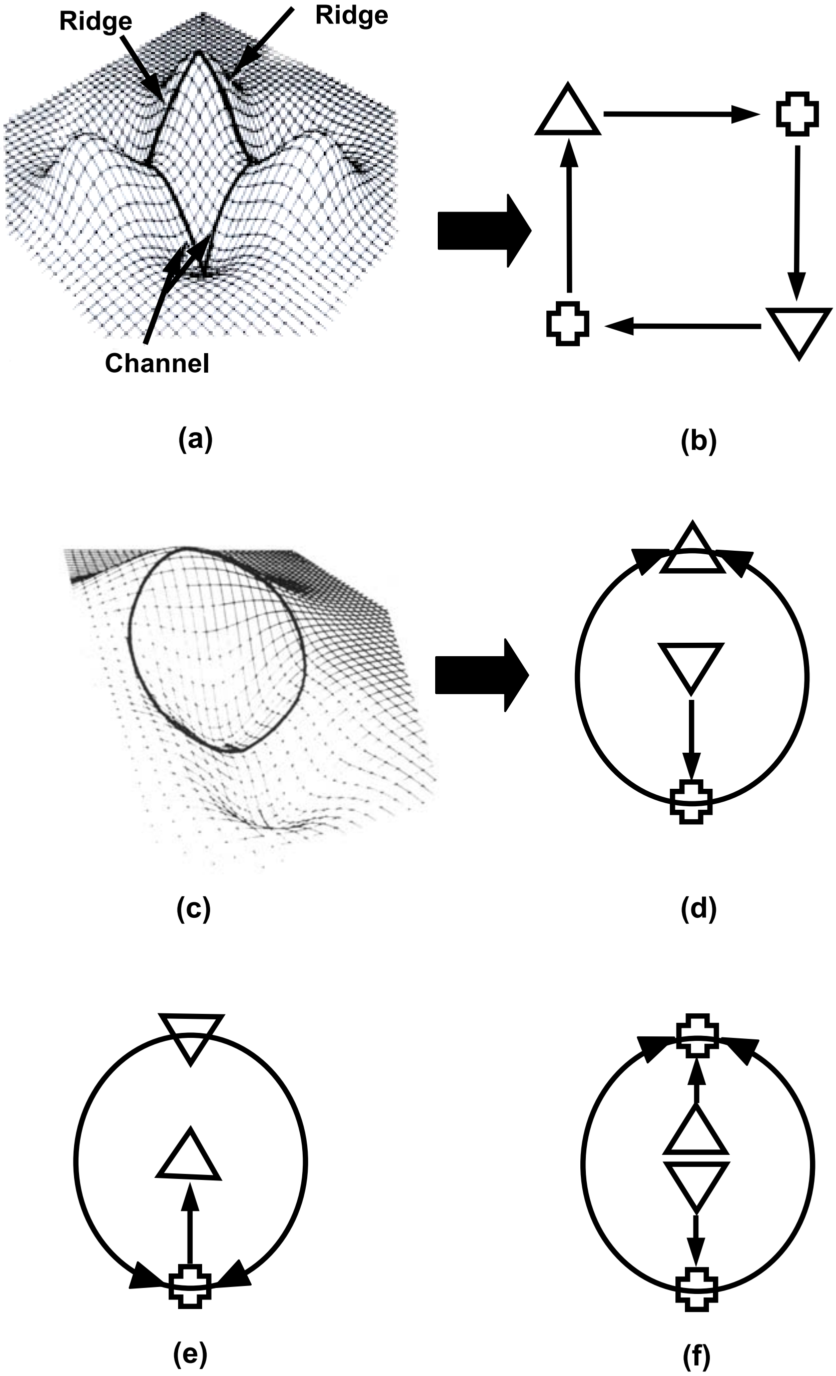


Figure 2.4 Two examples of slope districts - (a) and (c) and their Critical Point Configuration Graph - (b) and (d) respectively. (e) and (f) are the two other basic types of CPCG.

using differential equations and Morse Theory that the surface i.e., the CPCG, under reasonable assumption, contains four basic cycle types or slope districts (Fig. 2.4b,d,e,f). Nackman, however, did not propose how these slope districts could be conglomerated or paste together to form a single representation of the surface. This was perhaps one of the main reason for the lack of wide interest in CPCG (Recently revived by Rosin, 1995; Scott, 1998). In addition, as can be seen in the slope district at lower right (Fig. 2.4f), a pass can connect to pass with no intervening peaks or pits, which violates the rules laid by the theory of Hills and Dales for the ridge lines and course lines. Pfaltz (1978) reported that it is easy to create such surfaces mathematically (Morse, 1964) but remained uncertain if they could be used for terrains.

2.3 Summary

In conclusion for this chapter, it has been found that a generic treatment is still required to promote the surface network for wide and indiscriminate use. The following issues need to be addressed:

- Due to their design requirements, current implementations of Surface Networks have been restricted to surfaces with fluvial features (i.e., must have ridges, channels, peaks, passes, and pits). However, a number of surfaces have biased topography such as in glaciated or karstic terrains or features may be absent e.g., flat surfaces. Takahashi (1996) believes that the biased surfaces are cases of degenerate critical points and the presence of degenerate points leads to the violation of Euler criterion or Mountaineer's equation. His approach for handling degenerate points has been discussed in the next chapter.
- Points on the surface are classified as important points (pits, passes, and peaks) and lines (channels and ridges) based on the local slope or gradient around the points. This requirement restricted the implementation of surface networks on discrete surface data such as generated in social sciences. For this reason the triangulated irregular network (TIN) data structure will have to be used in these cases

although unlike surface network it would not provide any insights into the structure of the surface. However, it is also important to note that the approximation uncertainty will also usually be higher with the use of surface networks for discrete data.

In general, the absence of a general model is perhaps the reason for the existence of the nebula of different forms of surface networks mentioned in the last section. An aim of this is to bring together these ideas and propose a more general model of surface networks, which could then be implemented for most surfaces.

Chapter 3
Extraction of Surface Networks

3.1 Introduction

The process of accurate and indiscriminate extraction of a surface network from its surface lies in the middle of the surface network model and its use in practise. Therefore, the extraction will set the potential usability of the surface network data structures. In fact, the original motivation of this PhD was the opportunity of new ideas in the automated extraction of surface networks. It is well known that the theoretical ideas are often not easy to be implemented in practical computing. There is generally some level of compromise between the accuracy and the processing efficiency. For example, in the case of surfaces, a discrete DEM is a preferable representation against a realistic polynomial representation because it is easier to manage and generate it, although the uncertainties with the discrete representation could be significant.

The methods of surface networks extraction have ranged from the simple- manual (Wolf, 1984) and triangulation (Takahashi, 1995) to the complex surface fitting (Pfaltz, 1976; Wood, 1998). The different methods were chosen depending upon the researchers' belief on the best way of extracting the critical points and lines. There have been many suggestions for detecting the critical points and lines of a surface. This chapter will describe the above-mentioned four works in some details as they represent the culmination of the most widely implemented ideas and were used specifically for surface networks.

3.2 Surface Network Extraction Methods

Most simply, the generation of a surface network involves two steps – (i) extraction of the critical points and (ii) connecting them with the critical lines. However, the methods used for these two steps are still far more satisfactory. Two main concerns in the automated generation of surface networks are scale dependency and subjective feature definitions.

The issue of scale dependency is a multi-faceted and intensively studied topic across the academics. Various definitions and classifications have been proposed for scale and a number of books are dedicated in computer science (Lindeberg, 1994), earth sciences (Quattrochi and Goodchild, 1996) and social sciences on the determination and effects of scale in the processing. The basic issue, which concerns us, is that the features, objects and information exist across a range of scales, whose arrangement may and may not be hierarchical. At any one instance, our computing routines can detect features that fit into the fixed search window (kernel etc.). Therefore, the feature extraction could only be valid for the current scale but not as a true (natural) representation of the surface. In order to detect the scale, there have been attempts to model surface as fractals (Fels and Matson, 1996; Emerson and Quattrochi, 2000), wavelets (Starck et. al, 1998) or a simple hierarchical subdivision of surface (Csillag, 1996).

Despite the various approaches, there is no formal proposal on the effects of scale on features or objects on the surface. We believe that a part of the reason for the limited success in scale detection is due to the misunderstanding of the structure of the surface. Most feature detection techniques suffer from the inability to perform perceptual organisation of the local features into a more meaningful global scene. This is largely due to the absence of any prior information about the feature content of a surface and infinite combinations of features possible on the unknown surface. For simplicity, most scale detection methods studied the "scale-space" based on the behaviour of points on the surface. A prominent approach (in computer vision) has been measuring the appearance and disappearance of points under varying scales ("Gaussian Blurring") in the hope of being able to detect the spatial extents of the features (Lindeberg, 1994). However, most surfaces and images have a mix of features of different topological dimensions. In other words, the structure of the surface or image is actually made up of features of different topological types. For instance, in the case of surface networks, both the critical point and lines are important but they belong to different topological classes. The current algorithms do not take into account that different topological objects are expressed differently under different scales. For instance, points tend to be lost more quickly compared to lines over decreasing scales (zooming out). The conceptual issues such as "What is scale" and "What is the right scale of the surface?" also need to be addressed (Montello and Golledge, 1998).

Numerous methods and models have been proposed to characterise the critical points and lines. There is no consensus on the feature extraction technique but methods and algorithms are becoming more sophisticated (complicated) and universally available. The success of the algorithms depends on the critical point model i.e., eight neighbour methods or surface fitting and its scale dependency. For instance, some methods extract features in certain surfaces better than in other surfaces and some methods extract features better over only certain scales.

The main stress of this work so far has been to understand the various techniques and to put more efforts on the perceptual organisation of the features, building a topologically consistent surface network in this case. The following part of this chapter will describe some prominent methods for the automated, except one, extraction of surface networks.

3.2.1 Manual Extraction

In author's view, Wolf (1984) was perhaps the only successful researcher in extracting a topologically consistent surface network. The reason for his success lies in a manual extraction of the surface network from contour maps. He picked the critical points from the contour map using a digitiser and established the topological relationships i.e., the ridges and channels, by visual inspection. Wolf (1984) did not describe his methodology of the digitisation and therefore it will not be possible to give details about it in this report. A description of his work would have been useful because he may have followed an optimised methodology to extract a consistent surface network.

3.2.2 Triangulation

3.2.2.1 Definitions and Methodology

Takahashi et. al (1995) proposed a modified version of the eight-neighbour method based detection of the critical points (Peucker and Douglas, 1975) for grid surfaces. The eight-neighbour method compares the height of a point, $p(i,j)$, with its eight neighbours in a 3 x 3 square surrounding p (Fig. 3.1) and classifies the point as a critical point based on the criteria in Table 3.1.

i –1 , *j* -1	*i* –1 , *j*	*i* –1 , *j* +1
i , *j* -1	*p (i,j)*	*i* , *j* +1
i +1 , *j* -1	*i* +1 , *j*	*i* +1 , *j* +1

Figure 3.1 Point $p(i,j)$ in a grid (data view) and its 8 surrounding neighbours.

peak	$\lvert\Delta_+\rvert > T_{peak}$	$\lvert\Delta_-\rvert = 0$	$N_c = 0$
pit	$\lvert\Delta_-\rvert > T_{pit}$	$\lvert\Delta_+\rvert = 0$	$N_c = 0$
pass	$\lvert\Delta_+\rvert + \lvert\Delta_-\rvert > T_{pass}$		$N_c = 4$

$\lvert\Delta_+\rvert$ The sum of all positive height differences between the point and its 8 neighbours

$\lvert\Delta_-\rvert$ The sum of all negative height differences between the point and its 8 neighbours

N_c The number of sign changes associated with the point

T_{eak} Threshold height for a point to be a peak.
T_{pit} Threshold height for a point to be a pit
T_{pass} Threshold height for a point to be a pass.

Table 3.1 Criteria for classification of critical points in the eight-neighbour method.

Takahashi et al. (1995) showed that the eight-neighbour method based detection is subjective to the value of the threshold and this ambiguity could cause the loss of the Euler Formula property also called the Mountaineer's Equation i.e., $\text{pits} - \text{passes} + \text{peaks} \neq 2$. He suggested that in order to satisfy the Euler formula the contour changes should be determined according to the neighbour heights and not according to the threshold. He suggested the use of the Delaunay triangulation (Guibas and Stolfi, 1985) to triangulate the 3 x 3 square, centered at p, and

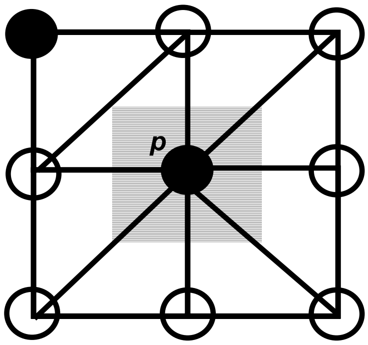


Figure 3.2 Point p in a grid (analytical view) and its 7 adjacent neighbours (hollow circles).

determine only the adjacent points (amongst the 8 surrounding neighbours) of p (Fig. 3.2). The point is then classified according to the criteria given in Table 3.2.

peak	$\lvert\Delta_+\rvert > 0$	$\lvert\Delta_-\rvert = 0$	$N_c = 0$
pit	$\lvert\Delta_-\rvert > 0$	$\lvert\Delta_+\rvert = 0$	$N_c = 0$
pass	$\lvert\Delta_+\rvert + \lvert\Delta_-\rvert > 0$		$N_c = 4$

Table 3.2 Criteria for the classification of non-degenerate critical points based on Delaunay triangulation.

However, in the case of degenerate passes (Fig. 3.3a) there will be more than 4 sign changes as three or more equi-height contours are merged. Takahashi derived that any degenerate pass can be decomposed into non-degenerate ones, m, where $m = (N_c - 2) / 2$ (Fig. 3.3d). By solving this equation, we can find out that the number of sign changes, N_c, at a degenerate pass will be equal to $2 + 2m$ ($m = 1,2,...$).

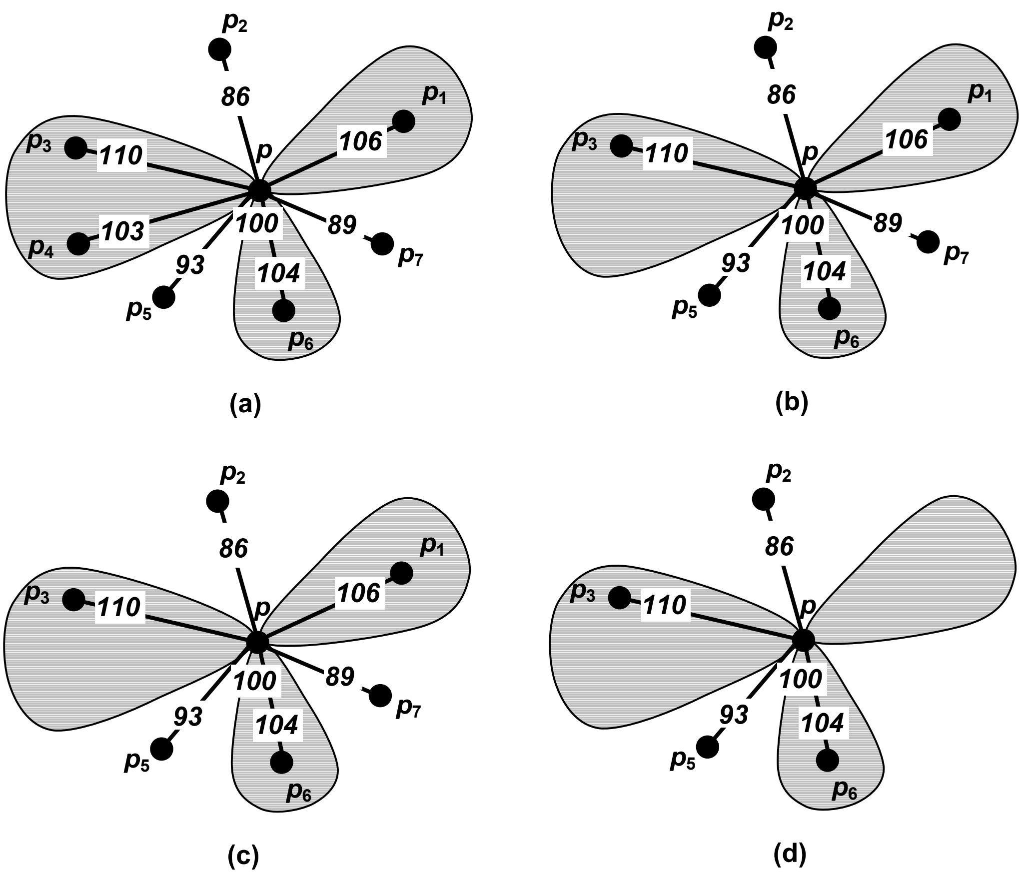


Figure 3.3 Decomposition of a degenerate pass (Modified from Takahashi et. al, 1995). Figure shows the neighbours and their heights. Higher neighbours are placed inside a grey region. (a) The original neighbour list, (b) the reduced neighbour list, (c) the list in the first turn of the loop in the algorithm, and (d) the final set of neighbours which will define the pass.

The algorithm to decompose a degenerate pass by Takahashi (1995) is unique and noteworthy. The steps are as follows :

(i) Generate a counter-clockwise (CCW) list of the adjacent neighbours of this pass, which in this case is $\{p_1, p_2, p_3, p_4, p_5, p_6, p_7\}$ (Fig. 3.3a).

(ii) Divide this list into an upper sequence, which has the higher neighbours, i.e, $\{p_1\}$,$\{p_3, p_4\}$ and $\{p_6\}$, and a lower sequence, which has the lower neighbours, i.e., $\{p_2\}$,$\{p_5\}$ and $\{p_7\}$. Reduce the neighbours list by selecting the highest neighbour from each upper sequence and the lowest neighbour from the lower sequence. For example, in the current example the original neighbours list is reduced to $\{p_2, p_3, p_5, p_6, p_7, p_1\}$ (Fig. 3.3.b) by removing p_4, because p_3

is higher in the sequence $\{p_3, p_4\}$. Note that if the list has more than one neighbour then the reduced list begins with a lower neighbour to ensure that the four alternating upper and lower neighbours at the pass are selected correctly. Also it can be seen from the reduced list that ther are 6 sign changes thereforre, the number of denegerate passes m is 2.

(iii) Put all the elements of the reduced list except the first two i.e., $\{p_5, p_6, p_7, p_1\}$, in a trailing list to further reduce the neighbours list.

(iv) Select the last four elements i.e., $\{p_5, p_6, p_7, p_1\}$, of the trailing list as *representative neighbours*. Remove the last two elements, which are $\{p_7, p_1\}$ in this case, of the *representative neighbours*, from the trailing list.

(v) Repeat steps (iii) – (iv) untill the trailing list is reduced to a lower and a upper neighbour of the pass, which in this case are $\{p_7, p_1\}$ and were easily achieved. The final neighbours list of the decomposed pass has the first two elements of the trailing list and the two elements remained after step (v) thus in this case the final neigbhours of p are $\{p_2, p_3, p_5, p_6\}$.

The methodology to connect the points is quite simple. It is based on the assumption that a ridge line is the line of steepest ascent from a pass while a channel is the line of steepest descent. Therefore, the ridge (channel) line is traced by moving to the highest (lowest) neighbour and repeating the tracing until a peak (pit) or the boundary is reached.

Takahashi et. al (1995) proposed that the above methodology would successfully extract a consistent surface network. However, we have some doubts, which will be shown in the next section.

3.2.2.2 Discussion

(a) Scale dependency: As mentioned earlier, features exist at various scales in a surface. The triangulation-based detection has a fixed scale of observation. It uses only the eight surrounding neighbours for the classification of the critical points. Takahashi (1996) was aware of this limitation and suggested referring to the scale-space theory (Witkin, 1983; Lindeberg, 1994). However, it is uncertain how the current method of triangulation can be extended to detect larger features.

(b) Limitations of feature classification:

- In order to avoid the inaccuracies related to the mathematical division of numbers, Takahashi et al. (1995) preferred the use of linear interpolation (Delaunay triangulation) to smooth surface (quadratic, cubic) fitting based methods to classify the points. See Wood (1996, 1998) for the disadvantages of the linear interpolation of heights for the classification of critical points and the extraction of surface networks.
- The ridge and channel lines are represented as the steepest lines of ascent and descent respectively from a pass, which again was debated by Wood (1996, 1998) as a proper method for feature identification.
- The decomposition of the degenerate passes is the unique aspect of the technique. However, the author suspects that the decomposition of the degenerate pass is rotation variant. For example, if we were to rotate the degenerate pass in Fig. 3.3a so that the neighbours lists starts from p_3 and not p_1 then the decomposed pass will have $\{p_5, p_6, p_7, p_1\}$ as the final neighbours. The author intends to take up this issue with Prof. Takahashi for confirmation before any further treatment.
- There is no proposal for the representation of junctions and bifurcations.

In the following section, the more sophisticated feature detection method, based on fitting a polynomial surface around a point, will be described. One of the main attractions of this method is its capability to perform multi-scale feature detection.

3.2.3 Polynomial Surface Fitting

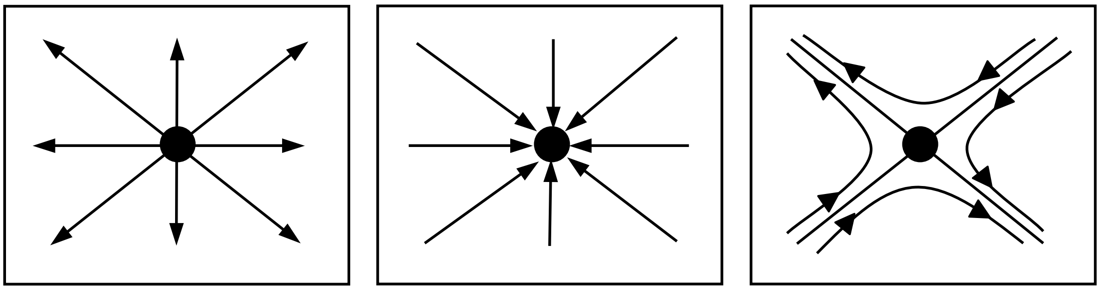
3.2.3.1 Definitions and Methodology

Recall from the last chapter that according to the Morse Theory, a point is a critical point of the surface if the local slope at the point is zero i.e., $\frac{\delta z}{\delta x} = 0$, $\frac{\delta z}{\delta y} = 0$. However, not all points that have zero slopes are critical points. In order to classify the locally flat areas into a peak or a pit or a pass, we have to know the local curvature using the second derivative of the height function at the candidate point. The local curvature can also be used to detect whether the candidate point is a ridge or channel. However, it is often advised to avoid the use of second derivative, as the second derivative tends to highlight the noise. The second derivative can be used to classify the critical points and lines in two ways. Firstly, the easier method is to compare the curvature along the three orthogonal components (see Table 3.3) (Wood, 1996). The components x and y are not necessarily parallel to the axes of the DEM, but are in the direction of maximum and minimum profile convexity. Secondly, the eigenvalues and

Feature Name	Derivative Expression	Description
Peak	$\frac{\delta^2 z}{\delta x^2} > 0, \frac{\delta^2 z}{\delta y^2} > 0$	Point that lies on a local convexity in all directions (all neighbours lower).
Ridge	$\frac{\delta^2 z}{\delta x^2} > 0, \frac{\delta^2 z}{\delta y^2} = 0$	Point that lies on a local convexity that is orthogonal to a line with no convexity/concavity.
Pass	$\frac{\delta^2 z}{\delta x^2} > 0, \frac{\delta^2 z}{\delta y^2} < 0$	Point that lies on a local convexity that is orthogonal to a local concavity.
Plane	$\frac{\delta^2 z}{\delta x^2} = 0, \frac{\delta^2 z}{\delta y^2} = 0$	Points that do not lie on any surface concavity or convexity.
Channel	$\frac{\delta^2 z}{\delta x^2} < 0, \frac{\delta^2 z}{\delta y^2} = 0$	Point that lies in a local concavity that is orthogonal to a line with no concavity/convexity.
Pit	$\frac{\delta^2 z}{\delta x^2} < 0, \frac{\delta^2 z}{\delta y^2} < 0$	Point that lies in a local concavity in all directions (all neighbours higher).

Table 3.3 Morphometric Features described by second derivatives (After Wood, 1996)

eigenvectors of the Hessian matrix (see section 2.2.3 in Chapter 2) can give information about the gradient flow at the critical point (Fig. 3.4). A critical point is a peak if the 2 real parts (R_1, R_2) of the eigenvalues of the Hessian matrix are positive indicating a gradient flow away from the critical point. A critical point is a pit if the 2 real parts of the eigenvalues of the Hessian matrix are negative indicating a gradient flow towards the critical point. In the case of the pass, the 2 real parts of the eigenvalues are of different signs. In addition, at a pass the eigenvector along the positive eigenvalue indicates the ridge line while the eigenvector along the negative eigenvalue marks the channel direction.



Peak: $R_1, R_2 > 0$ Pit: $R_1, R_2 < 0$ Pass: $R_1 < 0, R_2 > 0$

Figure 3.4 Critical points of the surface and the configuration of their eigenvalues and eigenvectors. R_1 and R_2 are the real parts of the eigenvalues.

In order to calculate the derivatives, the local surface around a critical point can be interpolated as a polynomial of the desired smoothness. For example, it could be modelled as a biquadratic function (Evans, 1980; Wood, 1996) or a bicubic function (Bajaj and Schikore, 1996). It is clear that the complex polynomials will provide a significantly generalised surface approximation and will take longer time to be solved. Complex polynomial will also characterise lesser extent of the surface because it requires larger neighbourhoods i.e., bigger kernels or filters, to reach a reliable solution.

For instance, the surface around a DEM grid cell can be represented as the following continuous quadratic function, made up of the sum of six terms (Wood, 1998):

$$z = ax^2 + by^2 + cxy + dx + ey + f$$

Various methods have been used to solve the surface polynomials for the coefficients such as simple combinations of neighbouring cells (Evans, 1980; Zvenburgen and Thorne, 1987) and matrix algebra (Wood, 1996). The properties of the continuous surface fitted on the discrete DEM values can now be derived analytically from the continuous function. For example, Evans (1980) defines steepest slope and aspect as follows:

$$slope = \arctan(\sqrt{d^2 + e^2})$$

$$aspect = \arctan(e / d) \text{ , where } (x,y) = (0,0)$$

Second order derivatives such as longitudinal and cross-sectional curvature can also be derived from the quadratic function (Wood, 1998).

A potential uncertainty with these surface measures is that they represent the value of the measure at a *point* at the centre of the quadratic function (Wood, 1998). This is appropriate for *point* measure such as solar incidence angles (used for biological applications). However, some properties such as the flow of water over a surface require some description of the surface away from the centre i.e. some properties are *areal* properties (Wood, 1998). Wood (1998) proposed that the extended flow directions (and other properties) away from the centre of the modelled surface can be measured by defining the quadratic function as a conic section. The conic section analysis can also help in classification of critical points and lines. The conic sections are elliptic, parabolic, hyperbolic, and planar (Kindle, 1950) (Fig. 3.5). The first three cases represent the critical points and lines, namely pits and peaks (elliptic), channels and ridges (parabolic) and passes (hyperbolic). The conic section analysis of the quadratic surface is especially useful in the cases when the centre of the critical point (line) is offset considerably from the centre of the area of interest (AOI). If the offset is significant, then the feature may be classified into the incorrect type. The benefit of using the conic section analysis is that the intersection between the semi-axes of the conic section and the region of interest can unambiguously determine the feature type and surface flow direction (Fig. 3.6). See Wood (1998) for the proof of this relation. This property thus can effectively handle the situation when the centre of the feature is offset from the centre of the AOI.

The procedure for connecting the critical points is more developed than the previous one because the information about the ridge and channel axes is also available (Wood, 1998; Wood and Rana, 2000). The steps are as follows:

(i) Identify the passes,
(ii) Move upwards in the direction of any ridge axes that fall within the AOI until a new grid is reached,
(iii) Recursively repeat (ii) until no higher cell is found,

(iv) Repeat steps (i) – (iii) but moving downwards along a channel axes.

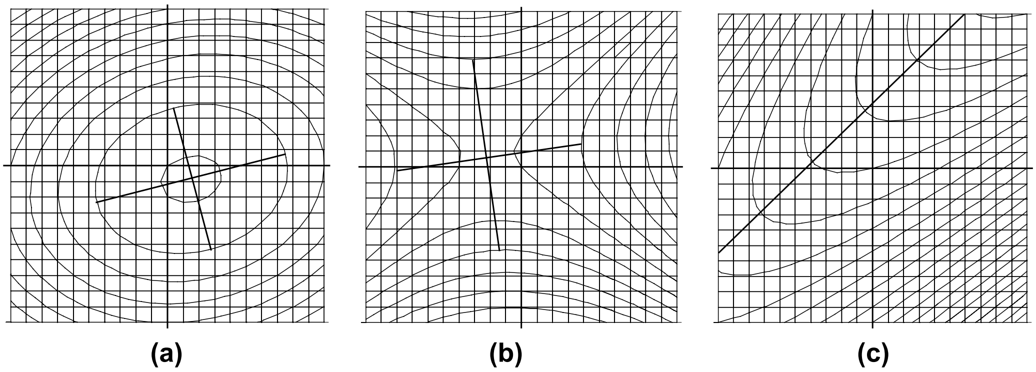


Figure 3.5 (a) Elliptic, (b) hyperbolic and (c) parabolic conic sections with their semi-axes identified (After Wood, 1998). A planar case is not considered here.

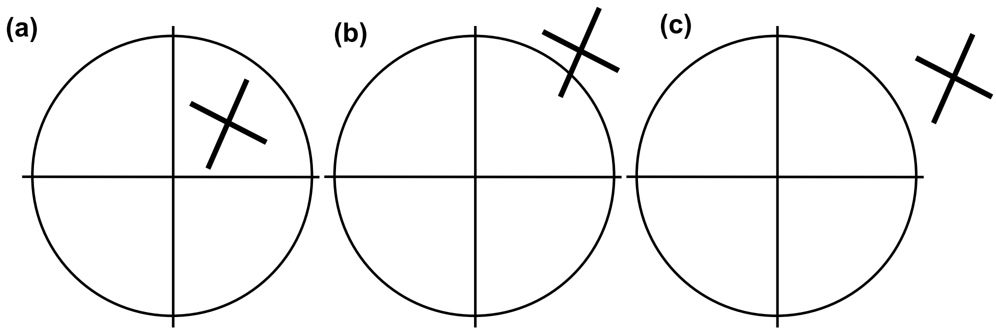


Figure 3.6 Three possible intersection cases between (circular) region of interest and conic section's semi-axes (After Wood, 1998). (a) Two axes intersect with region - pit, peak or pass, (b) one axis intersects with region - channel or ridge and (c) no intersection - planar.

3.2.3.2 Discussion

(i) Scale dependency: An advantage of the polynomial surface fitting based feature detection method over the earlier methods, is that it allows variable kernel size for feature detection. This allows the identification of features at various scales. However, there are no guidelines about the appropriate size of the kernels so the extraction is still scale-dependent. The lack of a proper scale analysis results into the loss of the Euler criterion (Wood, 1998). Also, fixed order of polynomial on varying scales may miss changes in surface complexity.

(ii) Limitation of feature classification: The feature extraction procedure does not perform any treatment of the degenerate points. Takahashi et al. (1995) showed that this is also a reason that the extracted surface network is inconsistent.

3.3 Next Research Aims

The following experiments are considered for further research in the extraction of surface networks:

(i) Scale-space detection:

- This will be explored if the triangulation technique for feature extraction can be modified to extract features at various scales.
- The behaviour of the critical points and lines in the scale-space will be compared to verify our view on the importance of topological dimensions in scale-space analysis.
- Since a theoretical treatment of the question "What is the right scale" seems open-ended, attempts will be made to achieve the answer based on empirical observations.

(ii) Feature definition:

- The accuracy of the triangulation- and polynomial surface-based feature detection methods will be compared for a variety of terrains.
- The rotation invariance of the methods will be confirmed.

Chapter 4

Generalisation of Surface Networks

4.1 Generalisation

In the case of generalisation in GI science, the abstraction and the uncertainties inherited in the abstraction (generally not recorded) of complex and large spatial information such as terrains, population, roads, is a major conern. According to Weibel and Dutton (1999), modern generalisation methods have basically two lineages namely from the generalisation in conventional cartography and the generalisation in digital systems (Fig. 4.1). "In conventional cartography, map generalisation is responsible for reducing complexity in a map in a scale reduction process, emphasising the essential while suppressing the unimportant, maintaining logical and unambiguous relations between map objects, and preserving aesthetic quality" (Weibel and Dutton, 1999). In other words, it involves techniques such as simplification, smoothing, aggregation and others. The generalisation process is somewhat one of post-processing. On the contrary, Weibel and Dutton (1999) believe that generalisation in digital systems inevitably starts at the stage of defining (or abstracting) a model for the detailed object (spatial information e.g., terrain). This stage called Object Generalisation is also a part of the generalisation in conventional cartography. After this stage, further reduction in the volume or precision of model data is often desired for compatibility with other data sets, easy portability across communication channels and faster analyses. This transformation is called Model Generalisation. The final stage in the generalisation in digital systems is of Cartographic Generalisation. The role of Cartographic Generalisation overlaps with that of Model

Generalisation in the sense that the result of both processes is the reduced data volume. However unlike Model Generalisation, Cartographic Generalisation has to take into consideration not only the scalar property but also vector properties, such as feature displacement directions (important to detect intersections after line simplifications) and aesthetic properties such as congestion and labelling.

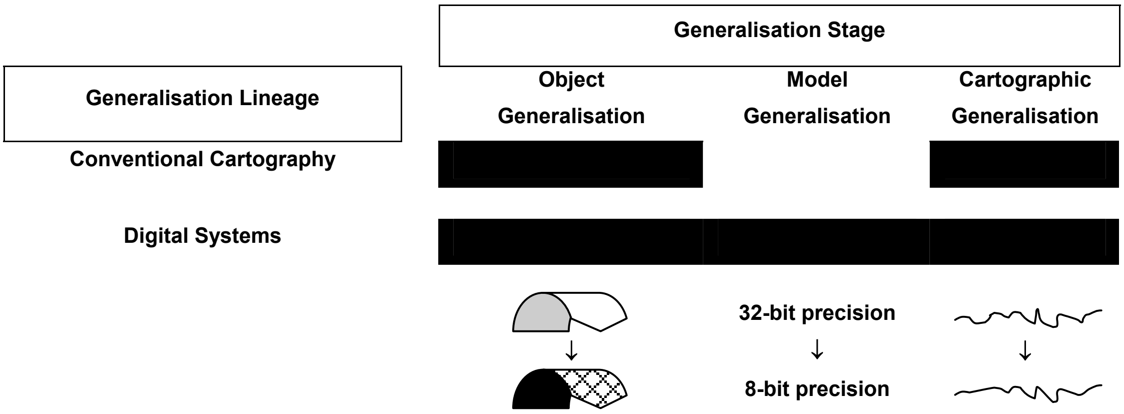


Figure 4.1 Generalisation Lineages and Stages in spatial data structures.

Müller (1991) compiled a list of requirements for generalisation methods. We suggest that these are also the benefits of the generalisation process. The main elements of this list are:

(a) Development of a model of the real world with an appropriate resolution and content,

(b) Efficient use of storage space and processing power,

(c) Development of a consistent and accurate database by removing spurious and redundant details

(d) Development of data and maps for various applications by suitably generalising details, and

(e) Optimisation of visualisation of data and maps.

We would like to add that the generalisation process could also help in understanding the structure of our spatial data especially if the data structure is based on a topological construction. Surface Networks will be a typical example of such a data structure. Later, evidence for this statement will be given. This knowledge about the structure would be subtly generated and destroyed in the generalisation processes listed above.

Based on the discussion above, it can be said that surface networks are an outcome of the Object Generalisation of surfaces. As the idea of surface network predates the digital age by centuries, they also serve as examples of Object Generalisation not being unique to generalisation in digital systems. As mentioned before, surface networks are a topological data structure therefore the Model- and Cartographic- Generalisation of surface networks have to be such that the resultant surface network should always be topologically consistent. For example, the surface network graph should be connected, all weights should be positive and such other properties of surface network mentioned before should be maintained. Pfaltz (1976) proposed the homomorphic contraction, a graph theoretic transformation, to prune sub-graphs of surface network graphs. Pfaltz viewed the generalisation as a Model Generalisation on the surface network i.e., to remove unimportant parts of surface network for clarity and efficiency. Wolf (1988) extended Pfaltz's idea in his PhD and proposed the idea of assigning importance to the critical points as weights and the contraction criteria. Wolf proposed homomorphic contractions as a way of performing Cartographic Generalisation of surface networks. Wolf showed (originally proposed by Mark, 1977) that the simplification of the topographic structure is a better alternative to produce simplified contour maps compared to the line simplification based generalisation of contour maps as the latter often produces contour intersections. We believe that the homomorphic contractions also provide us a unique and simple (compared to sophisticated numerical methods) technique of simulating surface evolution studies such as erosion modelling and others. This kind of simulation study lies somewhere in between Object Generalisation and Model Generalisation.

We believe that with the addition of the detailed geometry to the ridges and channels, traditional cartographic generalisation techniques also have a role to play in generalising surface networks. For example, it will generally be desired to simplify the channel and ridge paths using conventional line-simplifications methods. It is however unclear now what will be the requirements for the generalisation based on non-homomorphic contractions. A purely hypothetical outlook will be provided later in this chapter for the non-homomorphic contraction kind of generalisation of surface networks. A typical use of the non-homomorphic contractions will be given for the generalisation of unconstrained surface networks.

In the following sections, at first a brief discussion about the homomorphic contractions and non-homomorphic contractions will be given followed by the description of the methodology and results of some of our experiments carried so far.

4.2 Homomorphic Contraction of Surface Networks

Homomorphic contraction is an abstraction of a subgraph H of a graph G (V,E) to a single point, where V is the set of all vertices and E is the set of all edges The transformation can be formally stated as follows:

$G' = (V', E')$ is a simple homomorphic contraction of the graph $G = (V, E)$ if there exits a function $f_H: V \rightarrow V'$ such that:

(a) $f_H(v_i) = v'$ for all $v_i \in H$

$f_H(v_j) = v_j' \neq v'$ for all $v_j \notin H$

(b) $(v, w) \in E$ implies that $(f_H(v), f_H(w)) \in E'$ provided $f_H(v) \neq f_H(w)$.

However, due to the topological properties of surface networks not all homomorphic contractions can be applied to the surface networks. Wolf (1984) proposed two types of homomorphic contractions, which always result into a topologically consistent surface network. They are defined as follows:

(a) **$(y_o - z_o)$-contraction**:

Let,

- W = Surface Network,
- y_o = Pass with Peaks $R(y_o) = \{z_o, z\}$ and the difference in height along an adjacent ridge $h(y_o, z_o) <= h(y_i, z_o)$ for $i = 1,2, \ldots, n\text{-}1$ where $n =$ degree of the peak z_o.
- Set of adjacent passes to z_o $L(z_o) = \{y_o, y_1, y_2, \ldots, y_{n-1}\}$.

Then, (y_o, z_o)-contracted graph W' is the graph with the following properties:

- Vertex set $V(W') = V' = V - \{y_o, z_o\}$,
- Edge set $E(W') = E' = E + \{(y_1, z'), (y_2, z'),\ldots,(y_{n-1}, z')\}$, and
- Edge elevation drops:
    - $h(y_i, z') = h(y_i, z_o) - h(y_o, z_o) + h(y_o, z)$ for $i = 1,2,\ldots, n-1$.
    - $h(e') = h(e)$ for all other edges $e' \in E(W')$

This transformation, which contracts the subgraph $[y_o, z_o]$ and converts the original surface network onto a condensed one is called a (y_o, z_o)-contraction (Fig 4.2b). The contraction removes the peak z_o and its highest adjacent pass y_o together with all the critical lines incident with at least one of these critical lines. But this elimination causes the loss of two properties of surface networks, which are (a) the condensed subgraph $[P'_1, P'_2]$ is no longer connected (violation of rule P1 – see page 19-21) and (b) $od(y_i) = 1$ for $i = 1,2, \ldots, n-1$ (violation of rule P3 - see page 19-21). The topological consistency is restored by connecting loose passes y_i to z i.e., the edge set of W' contains the old edge set $E(W)$ and the new links (y_i, z'). The most important part of the contraction is the choice of y_o, which ensures that the elevation differences along the new links are always greater than zero. This idea actually originated from Mark (1977), when he proposed methods for the generalisation surface trees. Positive elevation differences are essential for the realisation of a topographic surface for instance a situation where a higher pass connects to a lower peak is unnatural.

(b) **$(x_o - y_o)$-contraction**:

A $(x_o - y_o)$-contraction can be similarly defined for the contraction of the subgraph $[x_o, y_o]$ (Fig 4.2c). The only difference is that the pass y_o is the lowest pass connected to the pit x_o. A surface network therefore can be condensed by repeated (y_o-z_o)-contraction and $(x_o - y_o)$-contraction until a desired level of simplicity is achieved.

In the case of surface networks, only the internal peaks and internal pits could be selected for contraction. The next section describes the basis of selecting the pits and peaks for contraction.

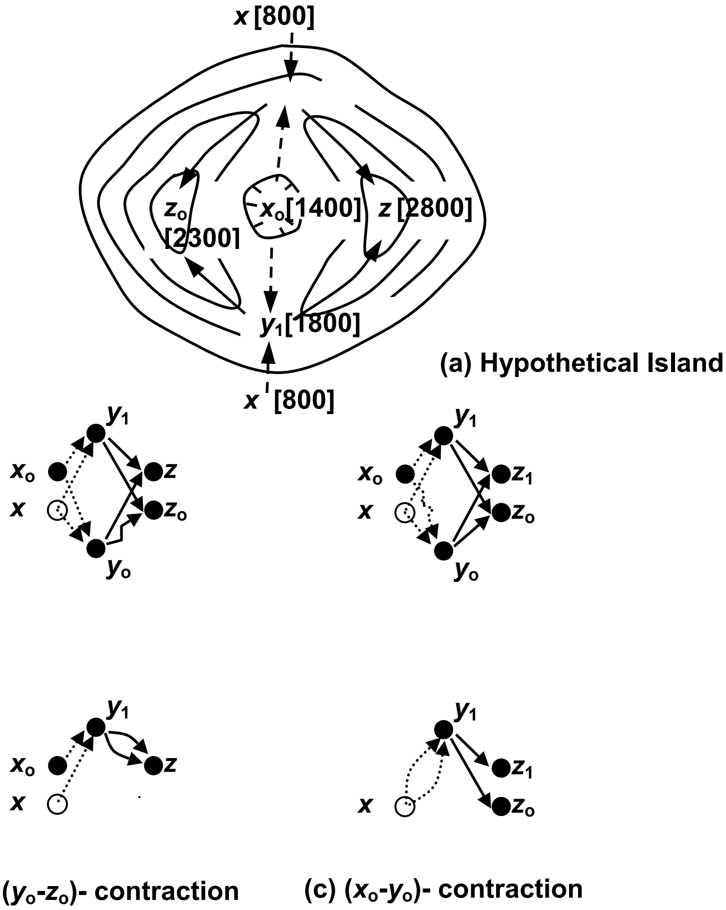


Figure 4.2 A hypothetical island and the (y_o–z_o)-contraction and (x_o–y_o)-contraction of its surface network. The numbers in the square brackets in (a) denote the height of the critical points and x is the surrounding pit

4.2.1 Criteria for Homomorphic Contraction

Although, the idea of homomorphic contraction was introduced by Pfaltz (1976) it was Mark (1977), Wolf (1985) and Rana (1998, 2000) who proposed various "importance measures" or weights, which could be used to select the peaks and pits for contraction. It is easy to realise that there could be many types of weights associated with the critical points. However, an ideal choice will depend upon the particular problem and the surface (Wolf, 1991). Mark (1977) and Wolf (1984) believed that any type of weights should be based on the elevation or in general on the value of the mapped property of the critical point because this would ensure a topologically consistent surface network after generalisation. The author however have observed that the importance of the critical point could be based on any measure which is suited to assess the importance of the critical point in the local or global neighbourhood. For instance, it could be the length of the edges, degree of the critical point, density of the local neighbourhood and many others. The critical step is to contract the edge with the least height difference (in case of terrains) or least mapped property. The following list of importance measures or weights is only a representative of many possible ways of assigning importance to the pits and peaks. Importance measure (i) – (v) are based on the suggestions of Mark (1977) and Wolf (1984) while importance measure (vi) – (viii) are based on Rana (2000) - work conducted during this research.

(i) **Height of the Peak and Pit.**

$\mathrm{w}\,(x_i) = |\, h(x_i)\,|$

$\mathrm{w}\,(z_k) = |\, h(z_k)\,|$

where (x_i) is a pit, (z_k) is a peak, h denotes height and w denotes weight. Height of the critical point is perhaps the simplest and most obvious weight that could be assigned to it (Mark, 1977).

(ii) **The maximum of the elevation differences between a peak or pit and all its adjacent passes.**

$w(x_i) = \max\ \{\ h(y_j) - h(x_i)\ \}$

$w(z_k) = \max\ \{\ h(z_k) - h(y_j)\ \}$

where $(x_i, y_j) \in E$ and $(z_k, y_j) \in E$. This measure can be used to remove peaks and pits ranked on the basis of the steepest ridge and channel linked to them.

(iii) **The minimum of the elevation differences between a peak or pit and all its adjacent passes.**

$w(x_i) = \min\ \{\ h(y_j) - h(x_i)\ \} = \min\ \{\ h(y_j)\ \}\ \text{-}\ h(x_i)$

$w(z_k) = \min\ \{\ h(z_k) - h(y_j)\ \} = \min\ \{\ h(z_k)\ \}\ \text{-}\ h(y_j)$

where $(x_i, y_j) \in E$ and $(z_k, y_j) \in E$. This measure can be used to remove peaks and pits ranked based on the shallowest ridge and channel linked to them.

(iv) **The sum of the elevation differences between a peak or pit and all its adjacent passes.**

$w(x_i) = \Sigma \{ h(y_j) - h(x_i) \}$

$w(z_k) = \Sigma \{ h(z_k) - h(y_j) \}$

where $(x_i, y_j) \in E$ and $(z_k, y_j) \in E$. This measure can be used to selectively remove pits and peaks with low number of crossings. However as can be seen this measure could be misleading because it will be biased by the heights of the points.

(v) **The sum of the elevation differences between a peak or pit and all its adjacent passes normalised by the degree of the peak or pit.**

$$w(x_i) = \frac{\Sigma \{h(y_j) - h(x_i)\}}{n(x_i)}$$

$$w(z_k) = \frac{\Sigma \{h(z_k) - h(y_j)\}}{n(z_k)}$$

where $(x_i, y_j) \in E$, $(z_k, y_j) \in E$ and n denotes the degree of the critical point. The idea behind this measure is same as in the last one but this one should remove the height dependency of the last measure. However, this is an unnecessarily long way of finding crossings. The degree of the peak or pit is perhaps more suited. Still, the normalised sum could prove to be useful for some other purpose.

(vi) **Degree of the Peak and Pit.**

$w(x_i) = n(x_i)$

$w(z_k) = n(z_k)$

The degree of the critical points i.e., the number of ridge and channels incident on the peak and pit respectively should be a quick and easy indicator of the crossings at these critical points.

(vii) **Maximum-, Minimum-, Sum-, and Normalised Sum- of the Length of the ridges and channels connected to the peak and pit.**

$w(x_i) = \max \{ \Lambda(y_j,x_i) \}$

$w(z_k) = \max \{ \Lambda(z_k,y_j) \}$

$w(x_i) = \min \{ \Lambda(y_j,x_i) \}$

$w(z_k) = \min \{ \Lambda(z_k,y_j) \}$

$w(x_i) = \Sigma \{ \Lambda(y_j,x_i) \}$

$w(z_k) = \Sigma \{ \Lambda(z_k,y_j) \}$

$$w(x_i) = \frac{\Sigma\Lambda(y_j, x_i)\}}{n(x_i)}$$

$$w(z_k) = \frac{\Sigma\Lambda(z_k, y_j)\}}{n(z_k)}$$

where $(x_i, y_j) \in E$, $(z_k, y_j) \in E$ and Λ denotes the length of the ridge or channel. The length measure is useful because it is perhaps a more realistic measure for the size (minor or major) ridges and channels. The assumption that a minor ridge will also generally have a small elevation difference is perhaps true in most surfaces but, in the presence of artificial or natural noise in the surface, this assumption could be misleading. More discussion will be given later in this chapter.

(viii) **Maximum-, Minimum-, Sum-, and Normalised Sum- of the Slope of the ridges and channels connected to the peak and pit.**

$w(x_i) = \max \{ \Delta(y_j,x_i) \}$

$w(z_k) = \max \{ \Delta(z_k,y_j)\}$

$w(x_i) = \min \{ \Delta(y_j,x_i) \}$

$w(z_k) = \min \{ \Delta(z_k,y_j) \}$

$w(x_i) = \Sigma \{ \Delta(y_j,x_i) \}$

$w(z_k) = \Sigma \{ \Delta(z_k,y_j) \}$

$$w(x_i) = \frac{\Sigma\Delta(y_j, x_i)\}}{n(x_i)}$$

$$w(z_k) = \frac{\Sigma\Delta(z_k, y_j)\}}{n(z_k)}$$

where $(x_i, y_j) \in E$, $(z_k, y_j) \in E$ and Δ denotes the slope of the ridge or channel. Slope of the ridge and channels is perhaps the ideal form of importance based on local neighbourhood as it includes both the height difference and the length of the ridges and channels.

As mentioned before there could be many ways of assigning weights. One particular aspect is that the weights mentioned above are all based on local neighbourhood but more global importance measures could provide more insights in characterising the surface.

We also feel that the weight is not the only way of selecting the peaks and pits for contraction. The sequential condensation of surface networks does not provide flexibility to the user to generate a desired topology and topography. Wolf (1989) experienced a typical limitation. He observed that the quality of condensed contour maps could be improved substantially if the step to eliminate a peak and its adjacent pass were shifted to a subsequent one. It is also apparent that vertex importance based selection criteria are insensitive to the ridge or channel structure at a peak or pit. This means that edges are solely selected for condensation, based on their weights and no consideration is given to the size or significance of the host structure (such as length of edges), which may not be suitable in some cases. Rana (2000) proposed the **User Defined Contraction (UDC),** in which a user can arbitrarily select an internal pit or peak for removal, which allows not only the flexibility desired above, but also the ability to create experimental surface networks.

In current ideas, there is a lack of suggestions to decide between equally weighted points. According to the Wolf (personal communication) either of the nodes can be selected arbitrarily as the other node(s) will be contracted in the next step(s) or perhaps an additional criterion for the order in which the nodes can be specified. It is easy to realise that this decision will have to be more sensible than an arbitrary one as the surface networks produced will be entirely different depending upon the choice. Another alternative is the use of a lexico-graphical basis for second ordering (Takahashi et. al, 1995). An example of such a situation is shown in Fig. 4.2, where the peak z_0 and the pit x_0 have two edges of equal weights. Wolf (1991) gave another example in his generalisation experiment. The use of this basis raises similar concerns like the previous basis therefore it is very essential for a researcher to be aware of this arbitrariness in the contraction purpose.

4.3 Non-Homomorphic contraction of Surface Networks

While the homomorphic contractions are useful for generalising the topology of surface networks, the addition of co-ordinates to the ridges and channels require that traditional line-simplification methods could also be used for the simplification of the geometry of the ridges and channels. This kind of simplification has been addressed extensively in a number of applications and the choice of a technique could be left to the user. The generalisation should of course produce topologically consistent surface network. Some topological properties will however be most vulnerable such as:

- **Planarity** – Line-simplification may cause ridges and channels to intersect (violation of P0).
- **Resemblance to original surface** – Too much simplification may cause channels to appear to cut across the surface. This would ultimately generalise to a purely topological state of the surface network, which reduce the visualisation potential of the surface network.

No attempts have been made in research so far on formalising the rules for this kind of generalisation of surface networks.

4.4 Generalisation Experiment

4.4.1 Methodology

Experimentation involved carrying out condensation of two surfaces both taken from original data by Wolf (1991) (Fig. 4.3a) and Wolf (1989) (Fig. 4.4a). Lack of an automated routine for the generation of a consistent surface network led to the use of a ready-made consistent network. These surface networks (Fig. 4.3b, Fig. 4.4b) for the surfaces were created manually i.e., by identifying the fundamental points and their relations manually. Fig. 4.3a is a hypothetical surface while Fig. 4.4a is a surface from an area in the Latschur Mountains of the Western Carinthia region in Austria. The original file format of the surface network was modified slightly into the following format.

Points

Point	Col. 1	Col. 2	Col. 3	Col. 3	Col. 4	Col. 5
Pit	*X*	ID	x	y	z	0 (if surrounding) or 1 (if internal)
Pass	*Y*	ID	-do-	-do-	-do-	
Peak	*Z*	ID	-do-	-do	-do	0 (if surrounding) or 1 (if internal)

Lines

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6
E	Y_{ID}	$X1_{ID}$	$X2_{ID}$	$Z1_{ID}$	$Z2_{ID}$

For example, part of the data for the Figure 4.4b surface network is as follows:

Y	y4	1.61	0.58	1150	
X	x5	0.77	0.45	1000	0
Z	z6	2.74	0.35	2200	1
E	y1	x1	x2	z1	z2

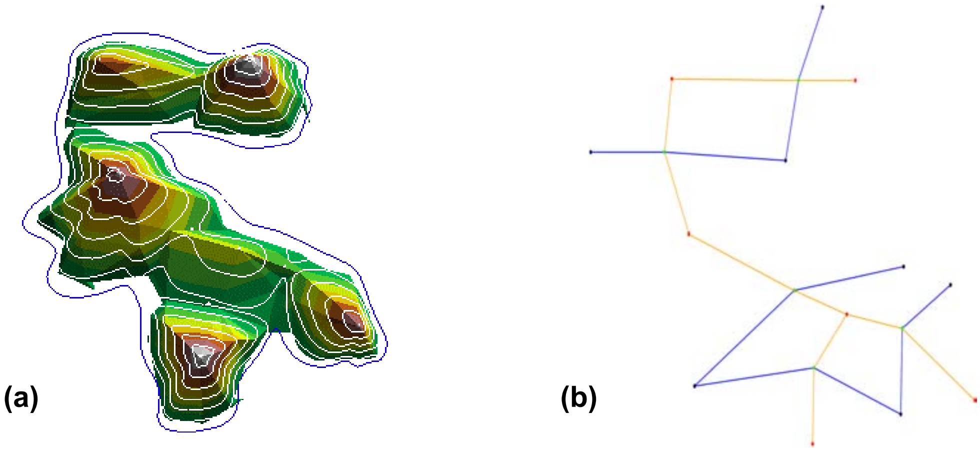


Figure 4.3 (a) Hypothetical topographic surface and its (b) surface network. Blue contour represents the surrounding pit.

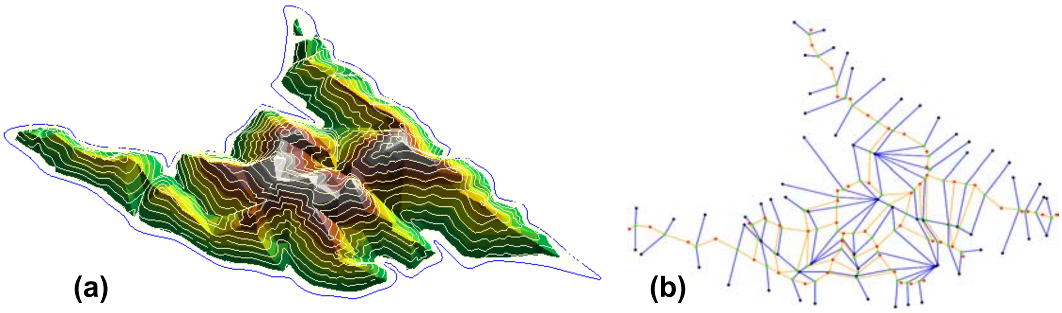


Figure 4.4 (a) Topography around the Latschur Mountains in the Carinthia Region, Austria and (b) its surface network.

4.4.2 Surface Topology Toolkit

The generalisations were carried out using an application, *Surface Topology Toolkit* (STT), developed by Sanjay Rana in Tcl/Tk. Tcl/Tk is becoming a popular language amongst GIS programmers (cdv by Jason Dykes, http://www.geog.le.ac.uk/jad7/cdv). The highlight of Tcl/Tk functions is the provision of dynamically manipulating the properties of graphical objects with ease and speed, which is particularly useful for cartographic and other visualisation applications. Owing to the Graphic User Interface (Figure 4.5) and UDC present in *STT* a user is able to achieve considerable improvement over the earlier methods for the generalisation and visualisation of surface networks (Wolf, 1991). The other main advantages of *STT* are as follows:

- *STT* informs the user of every contraction (except for continuous contractions) so that a selection can be made more intuitively.
- Users can generalise the topography by a combination of importance measure rather than a single one and can also arbitrarily select an internal pit or peak for contraction.
- Users have the flexibility to undo a contraction to observe the changes in results for better generalisation.

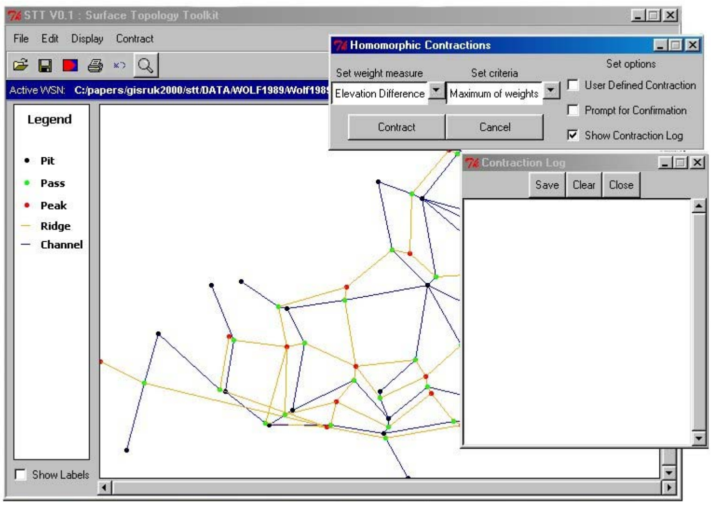


Figure 4.5 Graphical User Interface of the Surface Topology Toolkit application with the controls for the contractions.

An ArcView Avenue script has been implemented, which converts 2D surface networks into ArcView 3D-shape files so that they can be seen in 3D using ArcView 3D-Analyst extension.

This work has the following two experimental aims:

1. To compare the effectiveness of drop in elevation, edge length and valency weight measures.
2. To use UDC to generate artificial landform changes.

4.4.3 Results

Case 1: Effectiveness of drop in elevation, edge length and valency weight measure

As mentioned earlier the aim of using maximum and minimum edge weight criteria is the removal of peaks/pits based on respectively the steepest and shallowest ridges/channels linked to them. However, as would be expected a drop in elevation weight does not take into account the length of the edges, therefore it makes long edges vulnerable for condensation. For example, for the surface network shown in Fig. 4.6a, the next maximum drop in elevation weight criterion based condensation will remove the ridge $[y1, z2]$ (Fig. 4.6b) although it is longer, thus more important, than some of the other ridges in the surface. On the other hand, maximum edge length weight criterion based condensation selects to remove the ridge $[y1, z1]$ (Fig. 4.6c) and therefore is a more sensible measure. However, it is important to note that even after a better decision the ridge $[y1, z2]$ is still removed due to topology condensation rules, which proves the earlier stated proposal, that condensation solely based on weights, ignores the structure of ridge/channel networks.

Sum of edge weights and valency criteria are used to remove peaks/pits based on the ridge/channel crossings at them. The aim is to keep higher degree peaks or pits as they represent crossings of different ridge and channel lines and are therefore of great importance for the topography of the given area. A comparison of the condensation sequences based on the sum of drop in elevation weight criterion and valency weight criterion reveals that the later criterion identifies ridge/channel crossings more uniquely than the earlier criterion. Fig. 4.7 shows the situation in which of sum of edge weight criterion selects to remove the ridge [*y4, z5*] (Fig. 4.7b) although the peak *z5* has got the highest number of ridge crossings and is therefore a misleading condensation. On the other hand, valency weight criterion selects the ridge [*y5, z6*] (Figure 4.7c), which is closer to the expectation.

Case 2: Use of UDC to generate artificial landform changes
Study of landform evolution is a very useful topic of research in order to understand the geomorphic and tectonic phenomena in nature. Researchers use some form of landform models to simulate changes and predictions, but this often requires detailed mathematical analysis. As an alternative, this work proposes that UDC can be used to introduce similar changes more easily and quickly. An example of the generation of a NW-SE trending artificial valley in the Latschur surface network is shown in Fig. 4.8. This valley was achieved simply by merging minor channels in this area and the removal of the intersecting ridges along these channels. However, as its apparent, the changes are purely topological and one of the main advantages of other landform evolution models is their ability to regenerate the topography.

4.5 Regeneration of Surface

Even though the homomorphic contraction and line-simplifications are perhaps well established, according to the author the generalisation sequence is incomplete because there are no proposals for the general form of the surface inside and around the generalised part of the surface networks. In other words, we know about the connectivity of the critical points after a generalisation but we don't know how the ridges and channels should connect because we don't know what the "surface" looks like.

It is easy to realise that there could be infinite ways in which a river or ridge can meander but it is likely that most surfaces would behave in certain ways given a set of structural constraints. A common example of

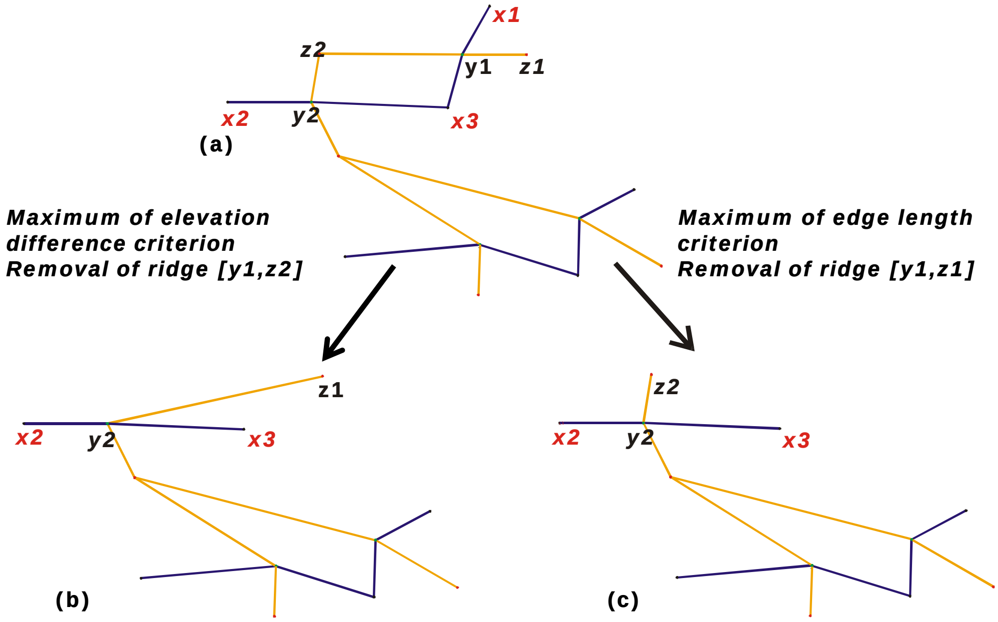


Figure 4.6 Comparison of the effectiveness for selection of points in the surface network (a) between maximum of elevation difference criterion (b) and maximum of edge length criterion and (c) Note that criterion (b) selects a long ridge due to its low drop in elevation (350).

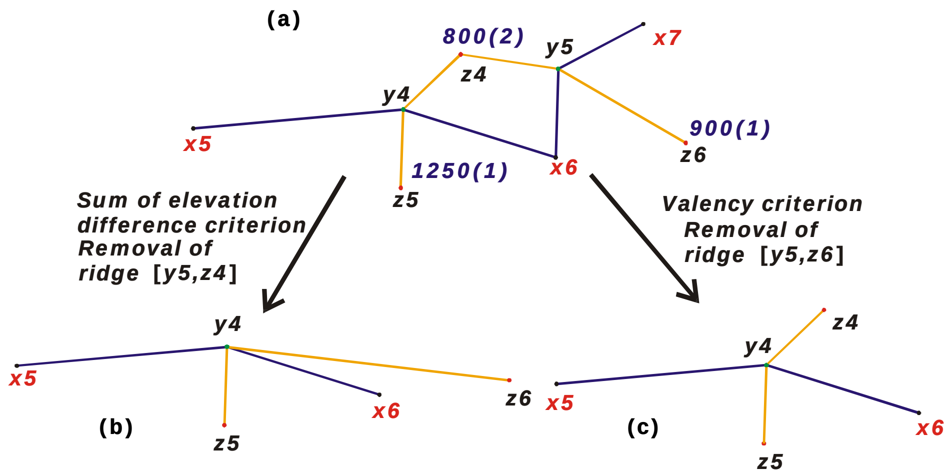


Figure 4.7 Comparison of the effectiveness for selection of points in a (a) surface network, between (b) sum of elevation difference criterion and (c) valency criterion, showing how criterion (b) can mislead about the ridge/channel crossings. Numbers at peaks in (a) are sum of elevation differences and their valencies (in parentheses).

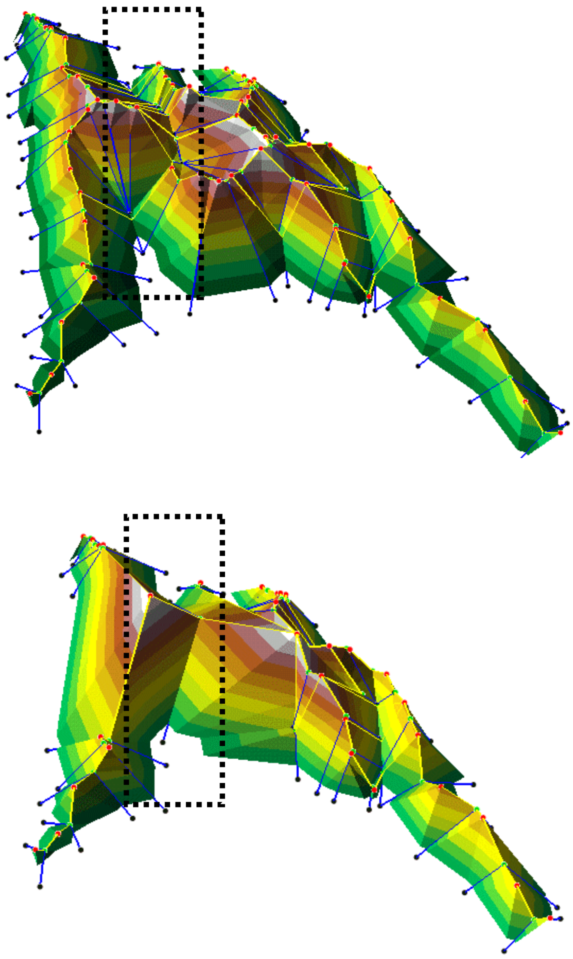


Figure 4.8 Generation of an artificial valley inside the dotted region of the Latschur surface network.

such an application is the erosion or aggradation modelling in geosciences. However, it is obvious that different surfaces will have different "evolution" models. For the time being, we would like to categorise broadly three kinds of approaches for regenerating the surface networks. These are arranged in increasing order of the likely complexity involved in the restoration method.

1. **Topological** – This kind of approach is purely the restoration of the topology of the surface network. The current methods of contraction already allow such regeneration. This will be the simplest and easiest possible method of regenerating a surface network.
2. **Artificial** – This method would involve some kind of artificial filling up of the generalised area with "surface like" details allowed under the topological constraints. Three methods are being explored namely Fractals, Region Merging (Takahashi and Kunii, 1994) and shape preservation (Bajaj and Schikore, 1997).
3. **Natural** – An ideal way of regenerating the generalised surface would be to simulate the generalisation as a form of natural surface process. A widely used example of such a modelling method is the terrain erosion modelling. These surface evolution models would require hypothesising solutions for morphological changes in the surfaces.

4.6 Discussion

Generalisation of Surface Networks particularly the homomorphic have an immense potential for future research. They could especially be useful for exploring the structure of a large surface. Some key areas, which will be addressed in the PhD, are:

- Detailed understanding of the property and effects of each of the condensation criterion.
- Further research on possible contraction criteria.
- Development of models for the regeneration of topography incorporating topological settings. A related issue is the refinement of surface networks (Bajaj and Schikore, 1997; Rosin, 1995).
- Formal methodologies for the non-homomorphic generalisation of surface networks.

Chapter 5

The Applications and Conclusions

5.1 Scope of Applications

Surface Network is a very "natural" representation of surfaces. It represents the surfaces in terms of the fundamental "surface elements" i.e., the peaks, pits, passes, ridges and channels. The use of the critical points and lines to represent the surfaces has a number of advantages, such as the following:

- It removes the subjectivity associated with the choice of legend (class intervals, colour scale etc.) to visualise the surfaces (Bajaj and Schikore, 1996). For example, there is often a certain level of uncertainty experienced while deciding the legend of a surface every time it is scaled or transformed. The critical points and lines, due to their unique positions in the surface, provide an intuitive understanding of the surfaces especially, which have complex structure such as dynamic maps (Rana, 2001a) and highly detailed surfaces (e.g., in flow topology by Helman and Hesselink, 1991). In essence, surface network is an intelligent data structure. In other words, if there were to be a measure of intelligence amongst the spatial data structures then surface networks will be much higher in the scale (Fig. 5.1). However, as the surface networks are a very coarse abstraction of the surface therefore they will score much higher on the scale of uncertainty in interpolating a surface based on them.
- The critical points and lines effectively act as landmarks on the surface thus they can be used to get a representative coverage of the entire surface. Intervisibility and viewshed analyses are particular examples

of modern applications, where such a property will be very useful (Rana, 2001b).

- According to the author, the homomorphic contraction of surface networks could be an easy and quick method of performing surface evolution processes, such as the erosion modelling in geology or morphing in computer graphics. Existing methods based on sophisticated numerical models are too computing intensive and require strong mathematical background.

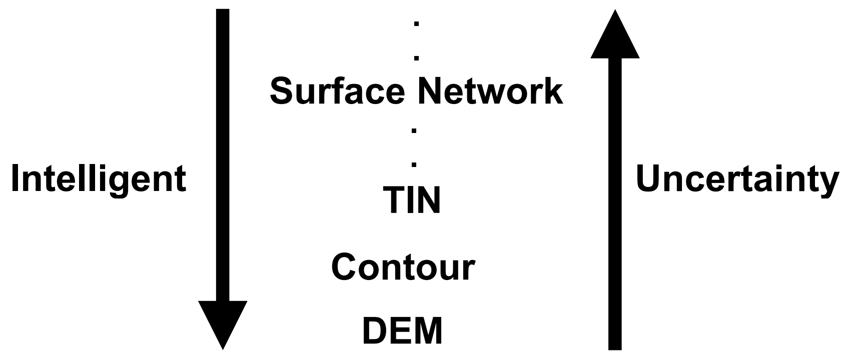


Figure 5.1 Comparison of the dumbness and uncertainty in spatial data structures.

In the following sections, some examples of ongoing experiments on the application of surface networks will be described.

5.1.1 Enhanced and Intuitive Visualisation

Complex surfaces such as dynamic maps and incised terrains have too many details to be sensibly interpreted by the viewer. Visualisation of dynamic maps is an actively discussed research topic in dynamic cartography (Shepard, 1995; Rana, 2001a). It has been debated whether the animation in dynamic maps is often a distraction rather than being of help in interpreting the map (Bertin, 1967; Dibiase et. al, 1992; McEachren, 1994). Bertin was perhaps right because although there have been many treatments of the elements of dynamic maps (dynamic variables) but there is still no guideline on the ways one should visualise a dynamic map.

As mentioned in the last section, surface networks provide a synoptic visualisation of the surfaces. This property could be useful in visualising dynamic maps, particularly used to represent continuous phenomena or fields such as weather maps. Fig. 5.2 shows an example of comparison between the traditional animation of weather patterns and surface network enhanced animation. It is clear that with the use of the surface network features it is easier to track the weather changes accurately. For example, the movement of the centre of the depression over the land can be more easily monitored. An interesting extension of this kind of visualisation will be to understand how the meteorological phenomena are linked to each other. For instance, the appearance, disappearance, and merging of the depressions could be observed by superimposing the framework of surface network. These ideas are still hypothetical and in the next stage, the author would like to pursue this field.

Few social science researchers have used the surface networks to visualise social phenomena such as urban settlement, commercial transactions and spatial flows (Warntz, 1966). We are exploring the possibility of work in these fields especially for their potentials to describe the structure of socio-economic phenomena.

5.1.2 Increased Efficiency

Critical points and lines are located at prime positions on a surface and therefore they can be used as the representative set over their local spatial neighbourhoods. In a number of analyses such as visibility and accessibility studies, an optimally located set of points is required to act as control to assess the significance of other points in the surface. For instance, a traditional problem in the visibility analysis has been the large processing time required assessing the visibility of points in large surfaces. There have been many proposals on decreasing the processing time. Most recently, based on their experiments with random points on the surface, O' Sullivan and Turner (2001) proposed that the critical points and lines are perhaps adequate to assess the Intervisibility of points on terrains. In this report, a preliminary experiment has been done to assess the potential of surface network to act as an observer framework for the Intervisibility of a terrain. In other words, instead of testing the visibility of each point against every other point, the visibility

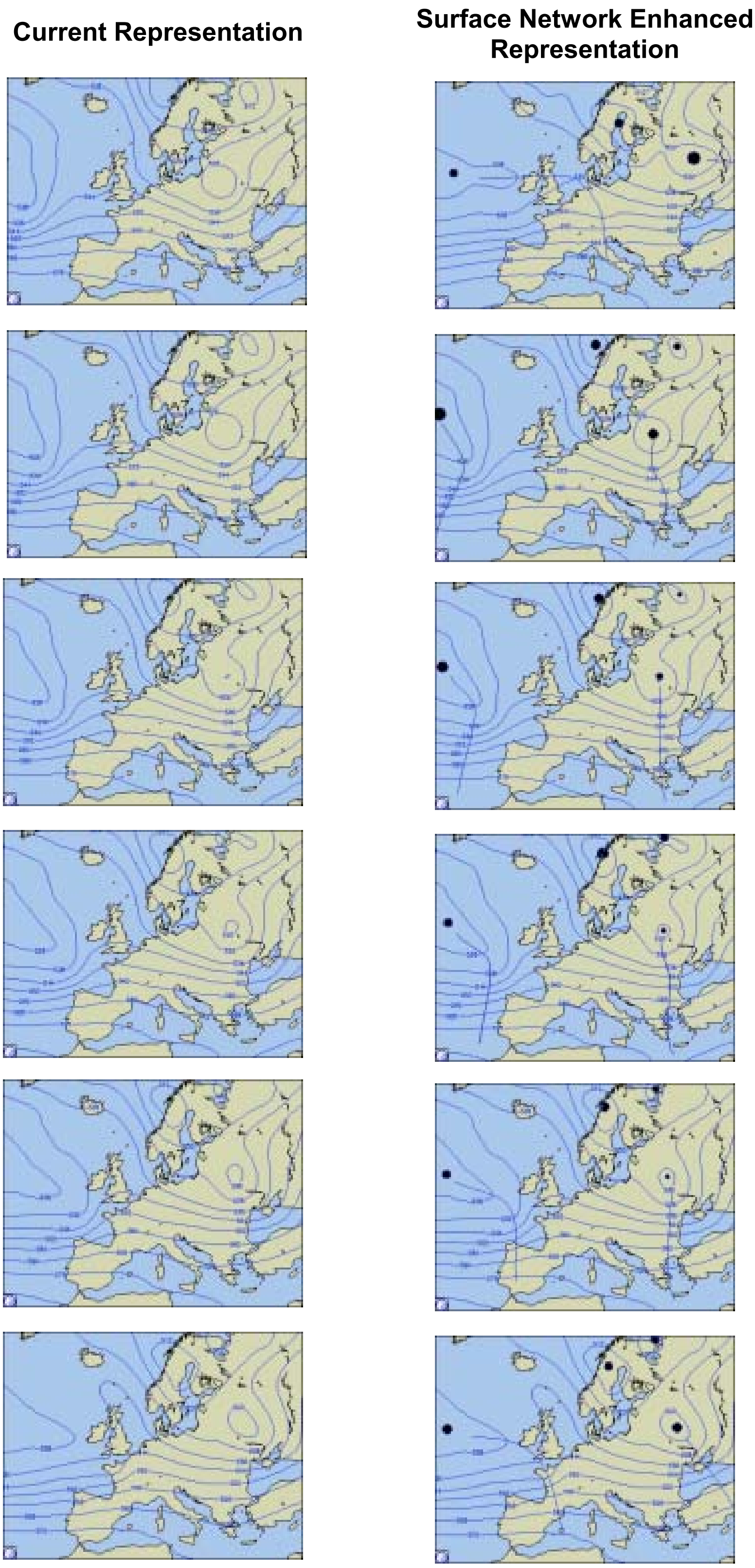


Figure 5.2 Comparison between the current and surface network enhanced visualisation of the dynamism in the geopotential height over Europe.

from a critical point and line was calculated (Fig. 5.3). Two observations have been made from a cursory comparison. Firstly, the overall pattern of the visibility of a point is very similar in both, with and without surface network based Intervisibility measurements. Secondly, as expected, the number of observers of a point in the surface network based Intervisibility is less than in the entire grid based calculation. However, is the low value significant? It is clear that a certain level of uncertainty has been introduced in the Intervisibility values. Can we quantify the uncertainty?

Overall, the potential of surface networks for Intervisibility studies still needs to be verified. We are developing this idea further with Dr. Y. H. Kim of Sheffield University to experiment with the use of surface networks for accessibility studies.

5.1.3 Simple Surface Evolution

Simulation of changes in the surfaces is a very attractive research area in many sciences. For example in geology, geomorphologists are interested in simulating erosion modelling and tectonic changes. In computer graphics, morphing is widely used technique in computer animation and effects. Social scientists are keen to experiment with different urban scenarios such as installation or removal of marketing town centres. Most of these operations require statistical and numerical methods for simulating the changes, which could be fairly complicated.

As suggested in Chapter 4 (section 4.5), the use of homomorphic contractions, both sequential and user-defined contractions, could be useful for performing simple surface evolution. However, it is clear that this idea needs to be thoroughly addressed before being advocated widely as a viable alternative. A critical issue in this idea, which needs to be resolved, is the regeneration of the surface after the "topological" generalisation.

5.2 Discussion and Conclusions

This report presents an overview of the previous works and author's research ideas on surface networks. The research on Surface Networks

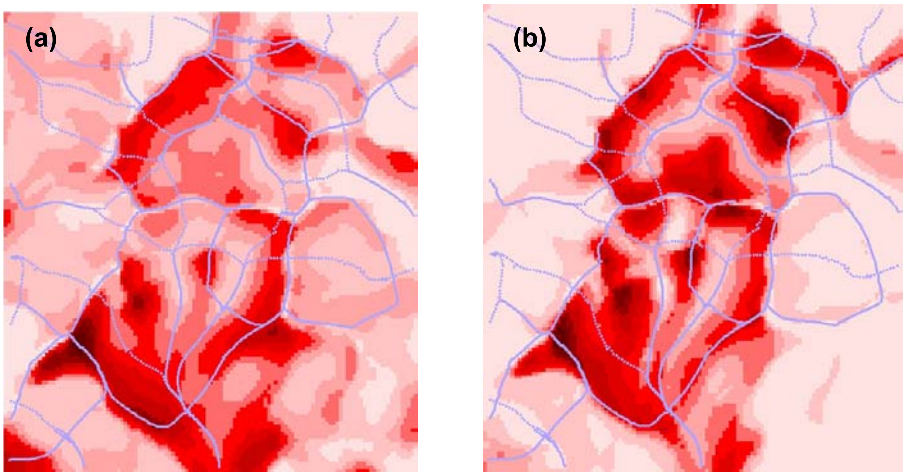


Figure 5.3 Intervisibility calculations in a part of Isle of Man, based on (a) each grid point (16335 points) and (b) critical points and lines (3975 points). Lighter to darker colour indicates increasing number of observers. The blue line is the surface network.

has four main focus areas namely, data structure model, automated extraction, generalisation, and applications. Since the start of research on parameterising a surface as a critical point-critical line framework in the 19th century, researches on this area has happened in different details in these four areas. We now clearly know and agree on the nature of the surface network components. However, many surfaces, especially natural terrains, pose problems in an automated extraction of surface networks. This is due to the inability of the automated routines to characterise the terrain intelligently into the critical points and lines. Some surfaces like natural terrains do not strictly follow the mathematically derived rules for surfaces. So what are the solutions? - the following two perhaps:

(a) Change the surface network design and rules to accommodate individual surfaces or

(b) Deliberately decompose the problematic parts of the surface somehow into what desired by surface network model.

The latter solution is probably easier and could be more natural (Of course depending upon the natural/realistic qualities of the decomposition process). Wolf (1990) and Takahashi et al. (1995) have shown examples of how junctions (bifurcations) and degenerate passes could be converted into surface network components. However, as discussed earlier, these decompositions have limitations. Therefore, designing a surface network model, which could be most adapted to most surfaces, is the first aim of this research. The automated extraction of a surface network is a related issue. The lack of a robust surface network model often creates problems in the automated extraction. However, the issues of a suitable feature definition model (e.g., how to detect whether a point is peak?), and scale dependency are internal to the automated extraction.

The methods and potentials of the homomorphic contractions of surface networks are practically unexplored. A large part of this research will focus on the issue of generalisation of surface networks.

The success of data structures comes down to their usability for practical applications. Surface Networks are used in different disciplines in different forms and names. A main aim of the author during the transfer period has been to establish contacts with researchers in various disciplines. It is hoped that the collaboration will ultimately help to produce a universal agreement on thoughts and perhaps even benefiting one another on some issues.

An issue, which is generally, ignored in most discussions on spatial data structures, is the uncertainty present in each abstraction of surface. It is perhaps difficult but not impossible to achieve a value for the approximation present in results derived from surface abstractions such as surface networks. We suggest that the approximation may vary according to the application and over the surface. For instance, the approximation in a visibility analysis may not be the same as in slope calculation. The approximation is likely to vary across the area of the surface because the approximation will depend on the density of the surface network in an area i.e., denser the surface network the lesser will be the approximation.

Finally, there is plenty of scope for new ideas in the research on surface networks, which have both intellectual and practical value.

Bibliography

Bajaj, C., and Schikore, D.R., 1996. Visualization of scalar topology for structural enhancement. Technical Report CSD-TR-96-006, Department of Computer Sciences, Purdue University.

Bertin, J., 1967. Semiologie Graphique. Den Hagg: Mouton.

Biasotti, S., Mortara, M., and Spagnuolo, M., 2000. Surface Compression and Reconstruction using Reeb Graphs and Shape Analysis. Spring Conference on Computer Graphics, Bratislava.

Cayley, A., 1859. On Contour and Slope Lines. The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, XVIII, 264-268.

Csillag, F., 1996. Quadtrees: Hierarchical Multiresolution Data Structures for Analysis of Digital Images. In: D. A. Quattrochi and M. F. Goodchild (Eds.), Scale in Remote Sensing and GIS, Lewis Publishers: Boca Raton, Fla., 247-272.

Dibiase, D., MacEchren, A.M., Krygier, J.B., Reeves, C., 1992. Animation and the role of map design in scientific visualization. Cartography and Geographic Information Systems, 19 (4), 201-204.

Emerson, C.W., and Quattrochi, D.A., 2000. Applications of fractal analytical techniques in the estimation of operational scale, Proceedings IAPRS, v. XXXIII, Amsterdam, Netherlands.

Evans, I.S., 1980. An integrated system of terrain analysis and slope mapping. Zeitschrift fur Geomorphologie, Suppl-Bd 36, 274-295.

Fels, J.E., and Matson, K.C., 1996. A Cognitively-based Approach for Hydrogeomorphic Land Classification using Digital Terrain Models. In: Proceedings 3rd International Conference/Workshop on Integrating GIS and Environmental Modelling, Santa Fe, New Mexico, USA.

Feuchtwanger, M., Poiker, T.K., 1987. The Surface Patchwork - an intelligent approach to terrain modelling. In: Proceedings 5th Annual North West Conference on Surveying and Mapping, Whistler, USA.

Fowler, R. J., Little, J.J, 1979. Automated extraction of irregular network digital terrain models. Computer Graphics 13, 199 - 207.

Helman, J.L., and Hesselink, L., 1991. Visualizing Vector Field Topology in Fluid Flows. IEEE Computer Graphics and Applications, 11(3), 36-46.

Kweon, I.S., and Kanade, T., 1994. Extracting topographic terrain features from elevation maps, 59 (2), 171-182.

Johnson, C., Burnett, M., and Dunbar, W., 1999. Crystallographic topology and its applications. In: P. Bourne and K. Watenpaugh (Eds.), Crystallographics Computing 7: Proceedings from the Macromolecular Crystallography Computing School, University Press.

Kindle, J.H., 1950. Theory and Problems of Plane and Solid Analytic Geometry, McGraw-Hill: New York.

Koenderink, J.J., and Doorn, A.J. van, 1979. The Structure of Two-dimensional Scalar Fields with Applications to Vision. Biological Cybernetics, 33, 151-158.

Lindeberg, T., 1994. Scale-space Theory in Computer Vision. Kluwer Academic Press: Dordrecht.

Mark, D.M., 1977. Topological Randomness of Geographic Surfaces of Geomorphic Surfaces. Technical Report #15, ONR Contract N00014-75-C0886.

Maxwell, J.C., 1870, On Hills and Dales. The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, XL, 421-427.

MacEchren, A.M., 1994. How Maps Work: Issues in Representations and Design, Guildford Press: New York.

Montello, D.G., and Golledge, R.G., 1998. Summary Report "Scale and Detail in the Cognition of Geographic Information. NCGIA, Santa Barbara, USA.

Morse, M., 1925. Relations between the Critical Points of a Real Function on n Independent Variables. Transactions of the American Mathematical Society, 27, 345-396.

Morse, M., 1964. The elimination of critical points of a non-degenerate function on a differentiable manifold. Journal Analyse Mathematics, 16, 257-316.

Morse, S.P., 1968. A Mathematical Model for the Analysis of Contour-Line data. Journal of the Association for Computing Machinery, 205 – 220.

Morse, S.P., 1969. Concepts of Use in Contour Map Processing. Communications of the Association for Computing Machinery, 147 – 152.

Müller, J-C., 1991. Generalisation of spatial databases. In D.J. Maguire, M.F. Goodchild, and D.W. Rhind (Eds.) Geographical Information Systems: Volume 1 - Principles and Applications, John Wiley and Sons Inc.: Harlow, Longman/New York, 457–475.

Nackman, L.R., 1984. Two-Dimensional Critical Point Configuration Graphs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 6 (4), 442-450.

O'Sullivan, D, and Turner, A, 2001. Visibility graphs and landscape visibility analysis. International Journal of Geographical Information Science, 15 (3), 221-237.

Palmer, B., 1984. Symbolic feature analysis and expert systems. In: Proceedings 2nd International Symposium on Spatial Data Handling, Zurich, Switzerland, 465 - 478.

Pfaltz, J.L., 1976. Surface Networks. Geographical Analysis 8 (2), 77-93.

Pfaltz, J.L., 1978. Surface Networks, An Analytic Tool for the study of Functional Surfaces. Final Report NSF Grant DCR-74-13353, Department of Applied Math and Computer Science, University of Virginia.

Quattrochi, D.A. and Goodchild, M.F., 1996. Scale in Remote Sensing and GIS. Lewis Publishers: Boca Raton, Fla.

Rana, S., 2000. Experiments on the generalisation and visualisation of surface networks. In: Proceedings GISRUK 2000, University of York, UK, 200-213.

Rana, S., 2001a. Role of vector information for augmenting visualisation of animated maps representing movements. Abstract accepted at GeoComputation 2001, University of Queensland, Brisbane, Australia.

Rana, S., 2001b. Optimising visibility analyses using critical points in the terrain. Abstract accepted at GeoComputation 2001, University of Queensland, Brisbane, Australia.

Reeb, G., 1946. Sur les points singuliers d'une forme de Pfaff completement intergrableou d'une fonction numerique. Computes Rendus Acad. Sciences Paris, 222, 847-849.

Reech, M., 1858. Proprieté générale des surfaces fermées. Ecole. Polytech J., 37, 169-178.

Rosin, P.L., 1995. Early Image Representations by Slope Districts, Journal of Visual Communication and Image Representation, 6 (3), 228 – 243.

Scott, P.J., 1998. Foundations of topological characterization of surface texture, International Journal of Machine Tools Manufacturing, 38 (5-6), 559- 566.

Shepherd, I.D.H., 1995. Putting time on the map: Dynamic displays in data visualization and GIS. In: P.F. Fisher (Ed.), Innovations in GIS 2, Taylor & Francis: London, 169-187.

Shinagawa, Y., Kunii, T.L., and Kergosien, Y.L., 1991. Surface Coding Based on Morse Theory, IEEE Computer Graphics and Applications, 11 (5), 66-91.

Speight, J.G., 1976. Numerical classification of landform elements from air photo data. Zeitschrift fur Geomorphologie Supplementband 25, 154 - 168.

Stark, J.L., Murtagh, F., and Bijaoui, A., 1998. Image Processing and Data Analysis – The Multiscale Approach, Cambridge University Press: Cambridge.

Takahashi, S., and Kunii, T.L., 1994. Manifold-based Multiple-viewpoint CAD. Computer-Aided Design, 26 (8), 622-631.

Takahashi, S., Ikeda, T., Shinagawa, Y., Kunii, T.L., and Ueda, M., 1995. Algorithms for Extracting Correct Critical Points and Constructing Topological Graphs from Discrete Geographical Elevation Data. The International Journal of the Eurographics Association, 14 (3), C-181- C-192.

Takahashi, S., 1996. Critical-point-based Modelling for Smooth Surfaces. PhD Thesis, Department of Information Science, University of Tokyo.

Tang, L, 1992. Automatic extraction of Specific Geomorphological Elements from contours. In: Proceedings 5th International Symposium on Spatial Data Handling, South Carolina, USA, 554 - 566.

Warntz, W., 1966. The Topology of a Socio-Economic Terrain and Spatial Flows. Papers of the Regional Science Association, 17, 47-61.

Weibel, R. and Dutton, G., 1999. Generalising spatial data and dealing with multiple representations. In: P.A. Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind (Eds.), Geographical Information Systems: Volume 1 – Principles and Technical Issues, John Wiley and Sons Inc.: New York, 125- 155.

Werner, C., 1988. Formal Analysis of Ridge and Channel Patterns in Maturely Eroded Terrain. Annals of the Association of American Geographers, 253-270.

Witkin, A.P., 1983. Scale-space Filtering. In: Proceedings 8th International on Artificial Intelligence, 1019-1022.

Wolf, G.W., 1984. A mathematical model of cartographic generalization. Geo-Processing, 2 (3), 271-286.

Wolf, G.W., 1990. Metric Surface Networks. In: Proceedings 4th International Symposium on Spatial Data Handling, Zurich, Switzerland, 844-856.

Wolf, G.W., 1991. A FORTRAN subroutine for cartographic generalization. Computers and Geosciences, 17 (10), 1359-1381.

Wolf, G.W., 1993. Data Structures for the topological characterisation of topographic surfaces. In: Proceedings Sixieme Colloque europeen de geographique theorique et quantitative, Theo Quant, Paris, France, 24-34.

Wood, J.D., 1996. The geomorphological characterisation of Digital Elevation Models. PhD Thesis, Department of Geography, University of Leicester.

Wood, J.D., 1998. Modelling the continuity of Surface form using Digital Elevation Models. In: Proceedings 8th International Symposium on Spatial Data Handling, Vancouver, Canada, 725-736.

Wood, J., and Rana, S., 2000. Construction of Weighted Surface Networks for the Representation and Analysis of Surface Topology, In: Proceedings GeoComputation 2000, University of Greenwich, UK.

Zevenbergen, L.W. and Thorne, C.R., 1987. Quantitative analysis of land surface topography. Earth Surface Processes and Landforms, 12, 47-56.