Surface Networks: Extension of the Topology and Extraction from Bilinear Surface Patches

B. Schneider

Department of Earth Sciences, University of Basel,
Bernoullistrasse 32, CH-4056 Basel, Switzerland.
Telephone: +41(0) 267 36 14
FAX: +41 (0) 267 36 13
Email: bernhard.schneider@unibas.ch

Abstract

Surface networks capture the topological relations between passes of a continuous surface, the paths of steepest descent and ascent starting at the passes, and the pits and peaks where the steepest paths end. This paper extends the topology of the network in three ways. Objects at the edge of the surface model are introduced. Horizontal areas may represent passes, pits, or peaks, and therefore must be detected and explicitly incorporated. And it is observed that there exist topographic configurations where paths of steepest descent and ascent intersect. These configurations and their consequence for the surface network topology are explained. To ensure consistency and completeness of the surface network, a zero-order continuous surface must be specified from the raster data prior to the extraction. The paper presents a method to derive the network elements from a surface composed of bilinear surface patches. The results are illustrated with two examples.

1. Introduction

Surface phenomena in the physical and socio-economic geosciences generally possess local minima and maxima, i.e., pits and peaks, as well as saddle points, i.e., passes. Each pass is starting point of two or more paths of steepest descent and ascent. The steepest paths end at pits and peaks, respectively, connecting the passes to the local extrema. The resulting graph of critical points and critical lines is termed *surface network* (Pfaltz 1976).

Surface networks are effective instruments to analyse topographic surfaces. Most obviously, surface networks yield the watersheds for all pits where each drainage area is bound by a sequence of ridge lines, passes, and peaks. Likewise, a ‘mountain’ – delimited by a sequence of valley lines, passes, and pits – is assigned to each peak. Furthermore, surface networks can be used to structure and characterise the (geo-)morphology of a surface (Warntz 1966; Werner 1994), to assist generalization of surface models (Wolf 1984; Rana and Wood 2000), and to detect spurious depressions allowing correction of these artefacts without alteration of the terrain data.

2. Continuity Constraints and Definitions

Fowler and Little (1979) propose deriving the network directly from raster data, without prior specification of a continuous surface. Critical points are detected with the method of Peucker and Douglas (1975) where the classification of a raster point depends on the relative elevations of the eight direct neighbours. The approach has a number of weaknesses that are related to the absence of a continuous surface:

- The locations of the critical points and of the critical line vertices are limited to the given raster points.
- Horizontal areas are not taken into account. Thus, the detection of critical points may be incomplete.
- The method detects too many passes, especially along crests and valley lines. As a result, the Euler formula is not satisfied (Takahashi et al. 1995).

Takahashi et al. (1995) reason that it is not possible to correctly extract critical points and to derive a consistent surface network solely from a set of discrete data points. Hence, a surface must be specified from the raster data prior to the surface network extraction.

On surfaces expressed as bivariate functions $z = f(x,y)$, pits, peaks, and passes are defined as points where the first derivatives in x and y are 0. Second derivatives are used to distinguish between the three types of critical points. For this reason, many authors (e.g., Wolf 1991, Rana and Wood 2000) request surfaces to be second-order continuous for the definition of the surface network elements. Unfortunately, real surfaces often have breaks in slope at many points and along many lines. Furthermore, their digital representations are not always continuously differentiable everywhere, if, for instance, the surface is represented by linear triangle facets of a TIN. As a consequence, it is necessary to re-formulate the definitions of the objects constituting the surface network without requesting second-order differentiability. (Henceforth, k-order differentiability is referred to as C_k-continuity.)

The following definitions of critical points and lines are valid for all globally C_0-continuos surfaces that are composed of piecewise C_1-continuos surface patches (De Floriani and Puppo 1992). They are consistent with the definitions given by Wolf (1991) and will serve as the basis for the discussion in this paper.

- A point *P* with elevation *z* on a C_0-continuos surface is called *pass* if
  a) there exists a circle around *P* with radius ***e*** along which the elevation changes at least twice from greater than *z* to lower than *z* (and, therefore, at least twice from lower than *z* to greater than *z)*;
  b) the above is true for all circles around *P* with radius *?* , *?* < ***e***, *?* > 0.
- A point *P* with elevation *z* on a C_0-continuos surface is called *pit* (*peak*) if
  a) there exists a circle around *P* with radius ***e*** along which the elevations are lower (higher) than *z*;
  b) the above is true for all circles around *P* with radius *?* , *?* < ***e***, *?* > 0.
- *Valley lines* are paths of steepest descent starting at passes. At each point *P* of the path of steepest descent where the surface is continuously differentiable, the tangent to the path of steepest descent coincides with the surface's aspect at *P*. If the surface is not continuously differentiable at *P*, *P* is at the border of two or more continuously differentiable surface patches. In this case, the path of steepest descent continues on the surface patch with the steepest slope at *P*. If all adjacent surface patches are higher than P in the vicinity of P, the path continues along the patch border with the steepest slope.
- *Ridge lines* are defined analogously to valley lines.

3. Basic Conceptual Model

In the geometric representation of the surface network, all points and lines are stored with their co-ordinates. Passes, pits, and peaks together with the valley and ridge lines form the frame of the network. Furthermore, a drainage area can be identified for each pit (Figure 1). Drainage areas are bound by sequences of ridge lines, passes, and peaks. This way, each pit is separated from its surrounding pits by ridge lines. Likewise, there exists a mountain for each peak, and there is always a valley line between two neighbouring peaks.

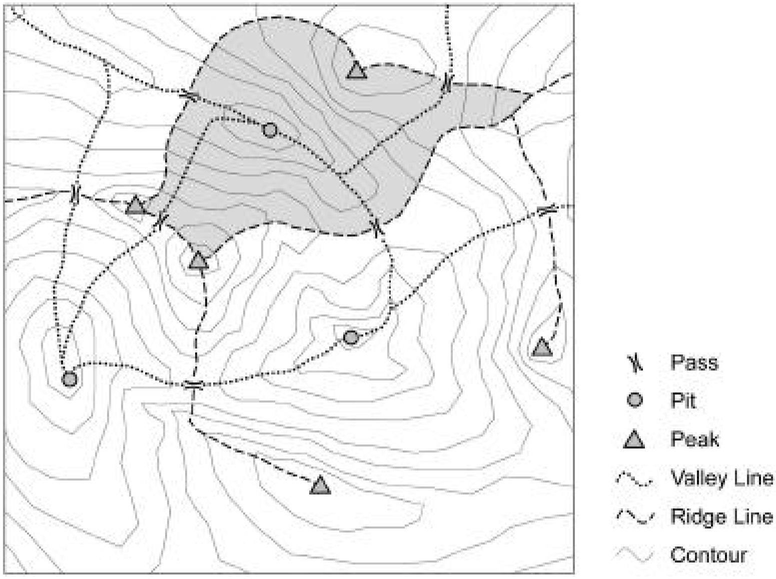

Figure 1. Geometric surface network of a hypothetical surface sketched with the help of contours. A selected drainage area is highlighted.

The *surface network graph* is the topological abstraction of the geometric surface network (Pfaltz 1976; Rana and Wood 2000) and denotes the topological relations between critical points and critical lines (Figure 2). Passes, pits, and peaks are the nodes, valley and ridge are directed edges of the graph. The co-ordinates of the points and line vertices are not important. Although they serve calculating the weights of the graph's nodes and edges in the Geometric Surface Network, they have no topological significance.

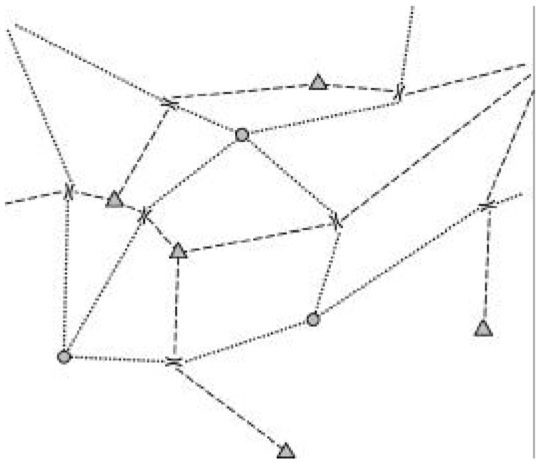

Figure 2. Topological surface network of the surface sketched in Figure 1.

4. Extending the Conceptual Model

4.1 Inclusion of Features at the Edge of the Area of Interest

It has been stated above that a drainage area can be delineated for each pit, and a mountain can be found for each peak. Figure 1 seems to disprove this statement because a border is lacking between the two mountains belonging to the two peaks in the lower right part of the area of interest. Since the valley line separating the two mountains does not exist, the pass where this valley line would start is obviously missing. After interpreting the contours, one expects the pass to be at the source of the valley that runs between the mountains and enters the area of interest just above the lower right corner of the area. However, because this pass is located outside the area of interest, it is not detected.

Topological consistency can be restored if topographic objects are introduced that owe their existence to the edge of the area of interest. Such objects are henceforth termed *edge objects* (Figures 3 and 4). Critical points on the edge are detected by, first, examining the profile of the topographic surface along the edge. All critical points manifest themselves as local minima and maxima of the profile. Second, the neighbourhood of the profile extreme points is analysed. If, for instance, a detected point is a local minimum not only of the profile, but of the surface then the point is an *edge pit*. Analogously, a local maximum of the profile is an *edge peak* if it is a local maximum .

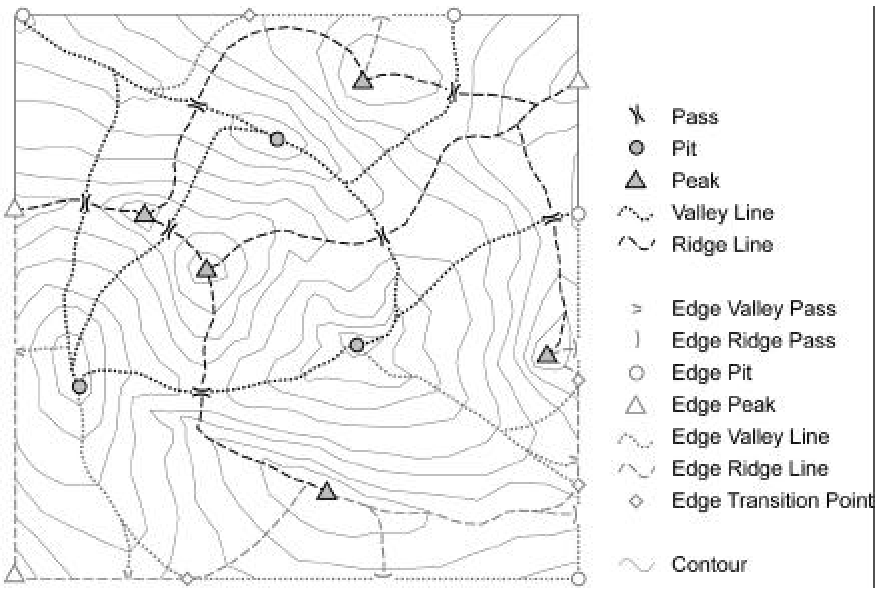

Figure 3. Extended geometric surface network.

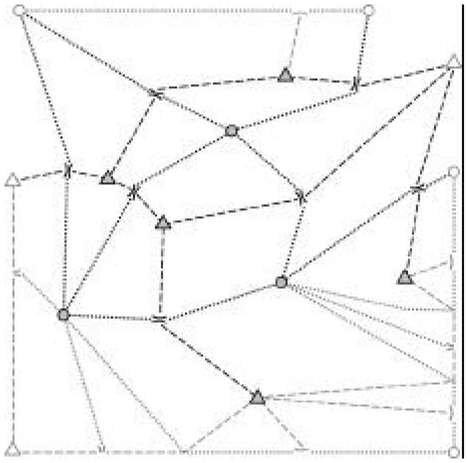

Figure 4. Extended topological surface network.

In both remaining cases where a local minimum (maximum) of the edge profile is not a local minimum (maximum) of the surface, the points are *edge passes*. In the case of a profile minimum, the point is an *edge valley pass* because a valley line enters the area of interest at that point. Since this valley line starts at the edge, it is termed *edge valley line*. The opposite valley line is located outside the area of interest and may be ignored. The two ridge lines starting at the pass, on the other hand, must be taken into account. These *edge ridge lines* initially follow the edge of the area of interest. They may deviate from the edge if the aspect vector of the surface along the edge changes from pointing into the area of interest to pointing outwards. Such locations where critical lines deviate from (or join, see below) the edge are termed *edge transition points*. After leaving the edge, the edge ridge lines follow the steepest path of ascent as ordinary ridge lines. Edge ridge lines end at peaks or at edge peaks.

In the second case where the local maximum of the profile is not a maximum of the surface, the point is an *edge ridge pass*. An edge ridge line enters the area of interest at this point. Also, two edge valley lines start at this pass and initially follow the edge.

All critical lines may join the edge at edge transition points. They may reach an edge pit or edge peak, or they may again deviate from the edge (Figure 3).

Valley lines and edge ridge lines may intersect at edge transition points. Previous authors (e.g., Wolf 1990) state that such intersections are prohibited. This suggests that such intersections are artefacts of the consideration of edge objects, and, thus, that this consideration is erroneous and invalid. This is, however, not the case. Instead, they may also occur without the edge of the area of interest being involved. The next section will describe these situations in more detail.

The definitions of pits and peaks given above remain valid if the exterior of the area of interest is considered void. Edge passes may be defined as special cases of regular passes:

- A point *P* with elevation *z* at the edge of a C_0-continuos surface model is called *edge valley pass* if
  a) there exists a circle around *P* with radius *e* where:
     i) the elevations of the intersections between circle and edge are greater than *z*;
     ii) there exists a point along the circle with elevaton lower than *z*;

b) the above is true for all circles around *P* with radius *?* , *?* < *e*, *?* > 0.

- *Edge ridge points* are defined analogously.

4.2 Interesection of Valley and Ridge Lines

Wood (1998; 2000) reports on the intersection of valley and ridge lines. Such intersections may occur in special topographic settings and must be explicitly addressed because they have fundamental implications for the topological surface network graph. Figure 5 illustrates a topographic situation leading to such an intersection. A crest and a valley run parallel to each other approximately south to north (if north is up). There is a pass on the crest and another one in the valley. One of the valley lines starting at the upper pass (left in Figure 5) initially runs down a gulch. At point *P*, the plan curvature along the path of steepest descent changes from a concave (i.e., the gulch) to a convex landform.

Figure 6 plots slope with respect to direction at *P*. There are two slope minima, that is, two directions with locally maximum downwards slope. Consequently, there are two options for the continuation of the path of steepest descent. The path continues along the steepest slope which leads towards the lower left of Figure 5. It may be argued that the choice is somewhat arbitrary and that the alternate path of steepest descent is valid too.

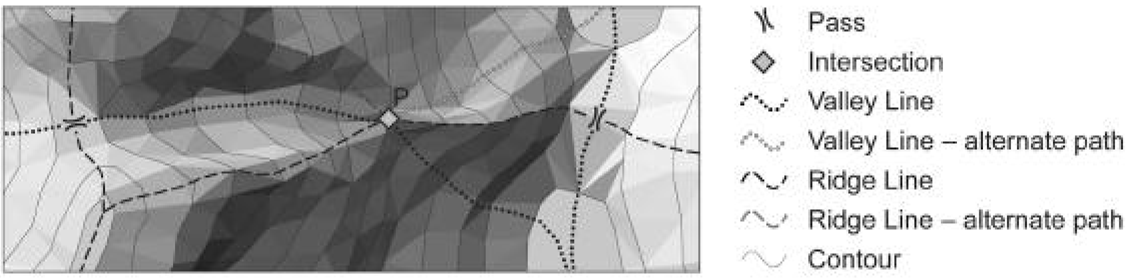

Figure 5. Extended Geometric Surface Network.

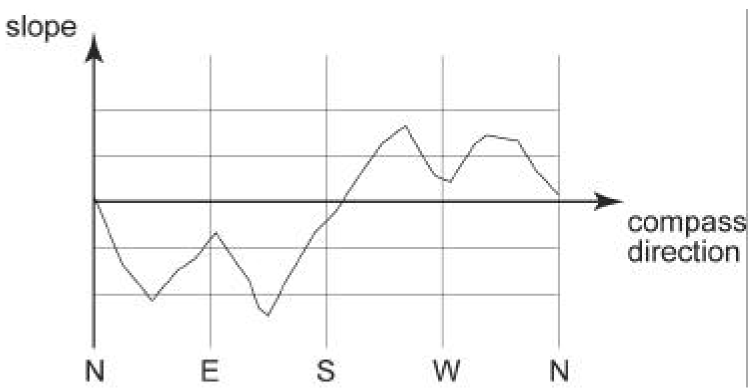

Figure 6. Directional slope profile around point *P* in Figure 5.

Notably, there is a locally maximum slope between the two minima belonging to a crest between the two possible paths of steepest descent. As Figure 5 illustrates, a ridge line may approach *P* along this crest. At *P*, this ridge line may continue towards the lower left or towards the upper left. The path towards the lower left happens to be steeper at *P*, and, thus, the valley line and the ridge line intersect at *P*.

Points where there exist more than one (locally) steepest slopes upwards and downwards along which a critical line may continue are labeled *bifurcations*. These points can be defined in accordance with the definitions of critical points:
A point *P* with elevation *z* on a C_0-continuos surface is called *bifurcation* if

a) *P* is not a pass nor a pit nor a peak;
b) *P* is not at the edge of a surface model;
c) there exists a circle around *P* with radius *e* along which the elevation has two or more minima *and* two or more maxima;
d) the above is true for all circles around *P* with radius *?* , *?* < *e*, *?* > 0.

Not every bifurcation is an intersection between valley and ridge lines. However, such intersection only occur at bifurcations.

Edge transition points discussed in the previous section are bifurcations at the edge because they may be the location of an intersection between an edge valley line and an edge ridge line. Edge transition points can be defined as follows:

- A point *P* with elevation *z* at the edge of a C_0-continuos surface model is called *edge transition point* if
  a) there exists a circle around *P* with radius *e* where:
    i) the elevations of one of the two intersections between circle and edge is greater than *z* while the other is lower;
    ii) the higher of the two intersections is a local minimum along the circle while the lower is a local maximum;
  b) the above is true for all circles around *P* with radius *?* , *?* < *e*, *?* > 0.

4.3 Horizontal Areas

Horizontal areas are frequent features in real surfaces. (One may argue that, in topographic surfaces, *exactly* horizontal areas other than water surfaces are rare). In digital surface representations, they occur even more often, caused by, for instance, the rounding of *z*-values of data points, or by triangular patches with all corner points on the same contour.

Figure 7 shows a surface composed by planar triangular facets specified from contours. The two triangles drawn with a horizontal line pattern together form a horizontal area. On two sides of this area (to the west and to the east), the adjacent triangles are lower, and on two sides (to the north and to the south), the neighbouring triangles are higher. Thus, the horizontal area acts as a pass. Likewise, horizontal areas may be pits, peaks, or bifurcations. Consequently, the topological significance of a topographic feature and its geometric representation must be separated.

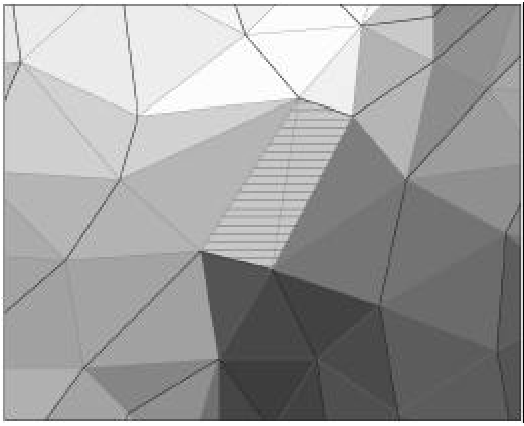

Figure 7. Horizontal pass area composed of two horizontal triangles.

The definitions of critical points given so far must be modified to comprise line and area objects. The circles with radius ***e*** are replaced by buffers around the objects with buffer distance ***e*** (i.e., by the boundary of the polygon that comprises all points with distance equal or less than ***e*** from the object). For instance, the definition of *pit* then reads:

- A point, horizontal line, or horizontal area object *P* with elevation *z* of a C_0-continuos surface is called *pit* if
  a) there exists a buffer around *P* with distance ***e*** from *P* along which the elevations are lower than *z*;
  b) the above is true for all buffers around *P* with distance *?* , *?* < ***e***, *?* > 0.

The interaction between horizontal areas and critical lines can be classified in two categories:

- A critical line leaves a horizontal area. This occurs if
  - the horizontal area is a pass and forms the critical lines' starting point;
  - the horizontal area is a generic horizontal area and the critical line passes through it.
- A critical line reaches a horizontal area. The line ends if the latter is a pit or a peak.

While the latter case is trivial, the former is not. When a critical line starts at a point object, the actual starting location naturally coincides with this point. When leaving a horizontal area, however, the exit point is not a given.

It may be presumed that the exit point lies on the border of the horizontal area. The search for the exit point of, for instance, a valley line may then be limited to the border sections where the immediate neighbourhood is lower. In Figure 7, there are two such sections. The exact location of the exit points, however, depends on the subjective interpretation of the horizontal area and on prior knowledge about the topography and the data portraying it. Possible approaches to locate the exit point of, for instance, a valley line are:

- In accordance with the definition of valley lines, the exit point is the point of the border where the immediate neighbourhood is steepest. This approach is not always practicable. In Figure 7, for instance, there is no such point because the adjacent triangles are planar and, therefore, slope is constant outside the border sections.

- The exit point is half way along the section neighbouring lower areas.

- The exit point is the point of the border closest to a hypothetical water source. If a valley line passes through a horizontal area, the water source is the point where the valley line reachies the horizontal area. In passes, the water source may be assumed to lie in the centre of the polygon (where centre needs to be appropriately defined, e.g., so that it lies within the area).

Finding the exit points of critical lines requests the application of heuristics. While this may seem a drawback at first, it must be considered a chance. The occurrence of horizontal areas in digital terrain models can not be avoided, and it is, therefore, mandatory to explicitly incorporate them in the construction of the surface network. As a result, however, it offers the opportunity to optimize the suitability of the results for a subsequent application.

5. Derivation from Surfaces Specified from Raster Data

5.1 Existing Approaches

Takahashi et al. (1995) propose triangulating regularly distributed DEM points prior to the surface network extraction. The triangles are generated by subdividing each grid cell with one of its diagonals. Of the two possible configurations in each grid cell, Takashiga et al. choose the one with the smaller angles between the planes defined by the two resulting triangles. The triangle-based surface allows using simple algorithms for identifying critical points and for tracing the critical lines. However, the subdivision of the cells into two planar triangles is based on a heuristic, and, thus, the extracted surface network is arbitrary to some degree.

Wood (1998) suggests specifying bi-variate quadratic surface patches at the raster points. These surface patches allow the identification of principle axes and, thus, of the directions of steepest descent and ascent. This information is used to extract the critical points and to trace the critical lines. The great benefit of bi-variate quadratic surfaces is that they can be fitted to windows of any size. This enables performing the analysis on a desired level of scale. However, the approach does not guarantee topological consistency of the extracted network. Depending on the shape of the surface and the scale specified, the method produces a number of non-connected sub-networks. Wood also reports on passes being falsely connected to passes because intermediary pits were not identified. Furthermore, the bi-variate quadratic surfaces tend to smooth the topographic surface with increasing window size. For instance, ‘sharp’ features (prominent breaklines) are not retained through the scales, and horizontal areas such as lakes shrink with increasing scale. It has not been investigated how this property of the bi-variate quadratic surfaces affects the extracted surface networks.

5.2 Choosing Bilinear Interpolation

As has been stated in the introduction, a C_0-continuous surface must be specified from the data. The bilinear interpolation is a straightforward and suitable choice. For each raster cell, a surface patch is specified with

$$z = a + bx + cy + dxy \tag{1}$$

where the coefficients *a*, *b*, *c*, and *d* are determined with the help the four corner points. The resulting surface patch is deterministic, that is, it is specified from the data directly, no intermediary data (e.g., vectors) need to be calculated, and no heuristic algorithm is involved. The global surface composed of the bilinear patches is, of course, characterised by apparent artefacts caused by the regular pattern of breaklines, i.e., by the borders between the cells along which the surface is not continuously differentiable. On the other hand, the bilinear interpolation scheme is conservative in the sense that the resulting surface does not overshoot or exhibit other artefacts known from higher-degree polynomial surfaces (Florinsky 2002; Schneider 2001). Furthermore, the bilinear scheme is computationally efficient, and extraction of topographic features is straightforward.

Figure 8. Bilinear interpolation. Right: bilinear surface patch. Left: globally C_0-continuous surface composed of bilinear surface patches.

The derivatives of Equation (1) with respect to *x* and *y* are:

$$z'_x = b + dy \quad (2)$$
$$z'_y = c + dx \quad (3)$$

Thus, slope does not change if one moves on the surface parallel to the *x*- or parallel to the *y*-axis. The profiles of the surface parallel to the co-ordinate axes are straight lines. This observation facilitates the detection of critical points and the tracing of critical lines.

5.3 Extraction of Critical Points

Bilinear surfaces do not have local maxima or minima. (There is only one point where the derivatives in *x* and *y* are both 0, and this point is a pass.) For this reason, the extrema of quadrilateral bilinear surface *patches* with straight borders parallel to the co-ordinate axes must be at their corner points. Thus, only grid points must be analysed to find pits and peaks. According to the above definitions, a grid point *T* is, for instance, a peak if there exists a neighbourhood around *T* in which the four adjacent bilinear surface patches are lower than *T* (Figure 9). This is the case if the corresponding tangential planes at *T* are lower than *T*. Such configurations can be identified by comparing the direct neighbours of *T* because each tangential plane at *T* is determined by the according straight lines from *T* to the direct neighbours (Figure 10). For example, if both, *Q* and *S* are lower than *T*, then the tangential plane – which is defined by the three points *QST* – is lower than *T* within cell C_1. Thus, it follows that *T* is a peak if the four direct neighbours *Q*, *S*, *U*, and *W* are lower than *T*. Pits are detected analogously.

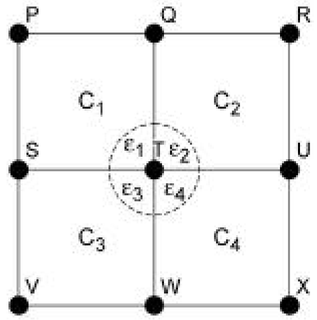

Figure 9. Scheme of the eight-point neighbourhood of raster point *T*.

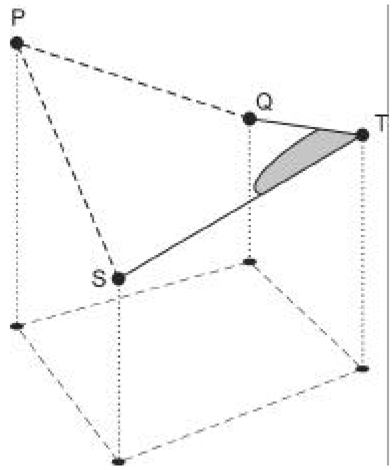

Figure 10. Tangent plane at corner point *T* of a square bilinear surface patch.

Passes may – in contrast to pits and peaks – occur not only at grid points but also within grid cells. If a grid point *T* is a pass, then its direct neighbours must be alternately higher and lower (e.g., *Q* and *W* are higher than *T*, *S* and *U* are lower than *T*; Figure 9). If a pass is located within a cell, then the corner points of this cell are alternately higher and lower. For instance, if *P* is higher than *S* and *Q*, and *T* is higher than *S* and *Q*, then there exists a pass within the bilinear surface patch. The surface at the pass is horizontal, i.e., the first derivatives in *x* and *y* are 0 which facilitates the calculation of the pass' exact co-ordinates.

5.4 Tracing Critical Lines

Critical lines are traced and constructed vertex by vertex. At each new vertex, the local configuration of the raster points (i.e., of the adjacent surface patches) is examined, and the next vertex calculated accordingly. A number of different situations can occur, all of which are listed in Table 1. Figure 11 illustrates a typical situation. The new vertex *B* of a valley line is found to be located on a cell edge (case ***b*** in Table 1). Examining the elevations of the corner points of the two adjacent cells 1 and 3, the edge is found to be a ravine, i.e., it is water-collecting (case ***ba***). Furthermore, the edge drops towards the raster point with elevation 42. Hence, the next vertex *C* is located on this raster point (case ***baa***). Vertex *C* becomes the new vertex (case ***a*** in Table 1), and again the elevations of all corner points of the four adjacent cells are examined. As in the previous situation, the cell edge towards the raster point with elevation 40 is a ravine (case ***aa***). However, after one third of the edge, cell

4 becomes lower than the cell edge. (The surface patch of cell 4 is horizontal along profile *p*.) As a result, the edge stops to be water-collecting after one third, and the path of steepest descent deviates from the edge to enter cell 4. This location on the edge is inserted as new vertex *D* (case ***aab***).

From *D* (case ***b***), the critical line continues through the cell (case ***bb***). Two possible methods to trace the path of steepest descent over the bilinear surface patch are:

- For each point of the surface patch, there exists an aspect vector. These vectors define a vector field. The path through a vector field is called streamline and can be calculated. In the case of a bilinear surface, the resulting curve is a hyperbolic function.
- Instead of calculating the entire path through the cell at once, it is approximated with small steps. At a point of the path, the aspect of the bilinear surface is calculated, and a small step of predefined length (e.g., a specific fraction of the cell size) is marked off along this direction. This process is repeated until the edge of the cell is reached, yielding vertex *E* in Figure 11 (case ***bbb***).

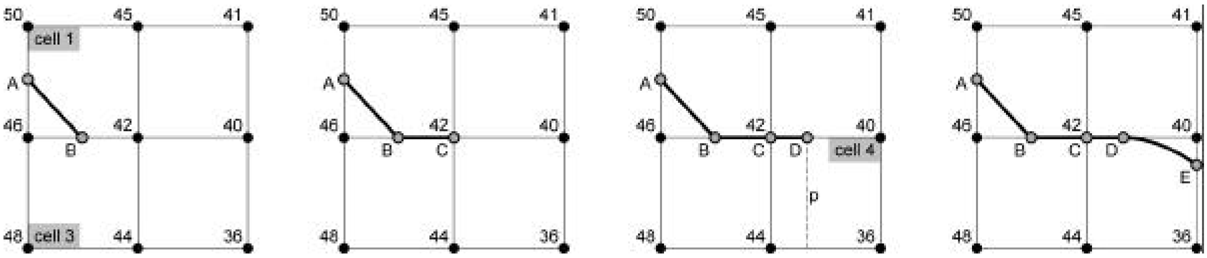

Figure 11. Four steps of the tracing of a valley line.

Table 1. Vertex locations, possible continuation of critical lines, and possible locations of next vertex.

location of new vertex	possible continuation	possible next vertex
a raster point	***aa*** along cell edge	***aaa*** next raster point along edge
		aab point on cell edge
	ab through cell interior	***aba*** diagonal raster point
		abb point on opposite cell edge
b point on cell edge	***ba*** along cell edge	***baa*** next raster point
		bab point on cell edge
	bb through cell interior	***bba*** diagonal raster point
		bbb point on opposite cell edge
c point inside cell	***ca*** any direction	***caa*** raster point
		cab point on cell edge

In addition to the above configurations, the newly calculated vertex of a critical line may lie on the edge of the area of interest. In this case, tracing the critical line is in accordance with the conception of model edges as discussed above. As in the ordinary case, critical lines are traced vertex by vertex, repeatedly examining the elevations of the relevant raster points, specifically considering the model edge.

Critical lines start at passes. At ordinary passes, i.e., at passes represented by point objects, two valley and two ridge lines begin. (Non-ordinary passes, i.e., passes represented by line and area objects, are discussed in the following section below.)

- If the pass coincides with a grid point, the steepest paths will continue along the cell edges (case ***aa*** in Table 1). The paths may reach the next grid point along the edge (case ***aaa***), or they may depart from the edge at some point between the two grid points (case ***aab***), depending on the two adjacent bilinear surface patches.
- If the pass is inside a grid cell (case ***c*** in Table 1), then the first segments of the steepest paths are straight and parallel to the cell diagonals until they reach the cell's edges (case ***cab***) or corner points (case ***caa***). (It can be proven that the paths of steepest descent and ascent on a bilinear surface starting at the surface's pass are straight and form an angle of 45° with the co-ordinate axes.)

5.5 Handling of Horizontal Areas

If two adjacent grid points have the same elevation, they constitute a horizontal edge. Multiple horizontal edges may form *horizontal edge groups* that may include horizontal cells (where all four corner points have the same elevation).

If a pass is represented by a horizontal edge group, it may be the start of more than two valley lines and two ridge lines. The number of steepest paths is computed by analysing all grid points directly neighbouring the horizontal edges. The exact location of each steepest path's start depends on the individual configurations of the transition from the horizontal edge group to its neighbourhood, and on the way this transition is morphologically interpreted. Likewise, a heuristic approach needs to be applied to trace steepest paths through horizontal regions.

6. Examples

The presented method has been applied to various synthetic and real surfaces of different characteristics and scales. The results of the extraction method are visually inspected.

The first of the two surfaces shown here is a synthetic surface generated by superimposing several sinus functions:

$$z = -\frac{5}{2}\sin\left(\frac{12px}{200}\right)\sin\left(\frac{8py}{200}\right) + r\sin\left(\frac{2px}{200}\right)\sin\left(\frac{3py}{200}\right) + \frac{10}{3}\sin\left(\sin\left(\frac{2px}{200}\right)\right) + \frac{x}{10} + \frac{y}{20} + 100 \qquad (4)$$

Values z are calculated for the range $0 = x = 200$, $0 = y = 200$. The cell size is 1 in both, the x- and y-direction. Figure 12 shows a hillshaded image, Figure 13 shows a 3D view of the surface and the derived network. In the two Figures – as in the remaining Figures – regular passes are drawn as green dots, edge pit passes are drawn in cyan-green, edge peak passes in yellowish green. Pits and valley lines are blue, peaks and ridge lines are red. Passes are drawn before pits and peaks which is why some passes are partly obscured by pits and peaks.

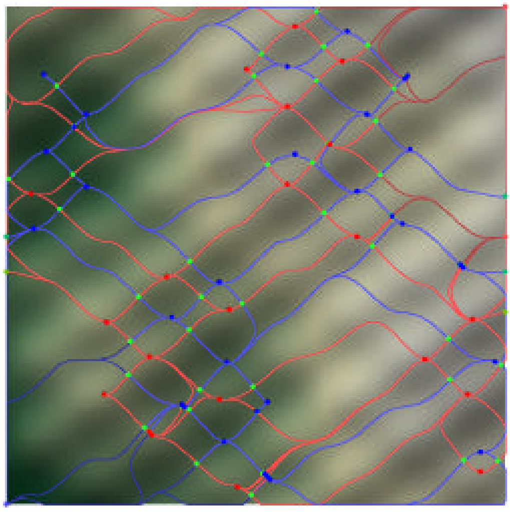

Figure 12. Surface network derived from the synthetic surface of Equation 4.

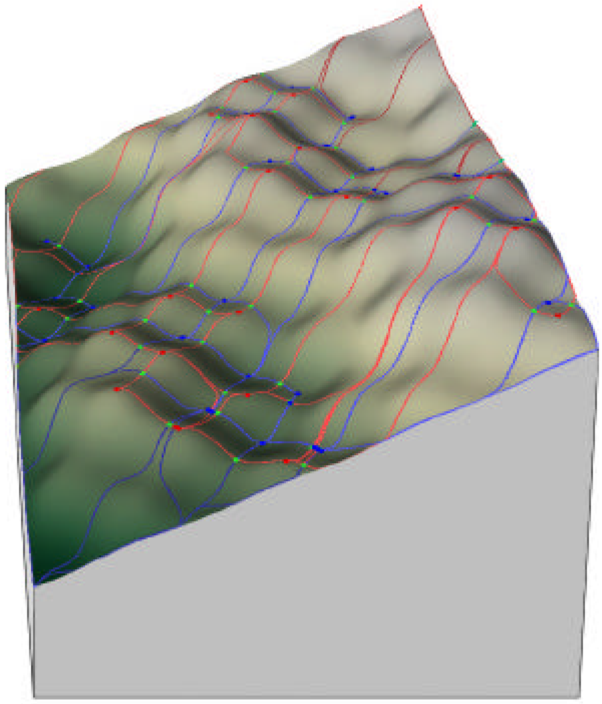

Figure 13. 3D view of the synthetic surface of Equation 4 and the according surface network.

The surface network is consistent in the sense that is connected (i.e., there are no disconnected sub-networks), and that all connections of critical points through critical lines are valid. All critical points are extracted, as far as visual inspection can determine, and for each critical point, an according area object can be discerned. At few locations, however, short chains of passes and pits or passes and peaks, respectively, occur.

Figures 14 and 15 show a test area located at the northern edge of the Alps in the eastern part of Switzerland, close to Druesberg, Canton of Schwyz. The elevations of the test area are between 1640 (upper left corner in Figure 14) and 2200 meters above sea level (lower right corner). The topography is characterised by a number of ridges and by a horizontal depression at the place of a dried out lake. The area has an extent of 1000 meters east-west and 920 meters north-south. The cell size is 11 meters, and the raster data are derived from 20-meter contours digitised from a 1:50’000 topographic map.

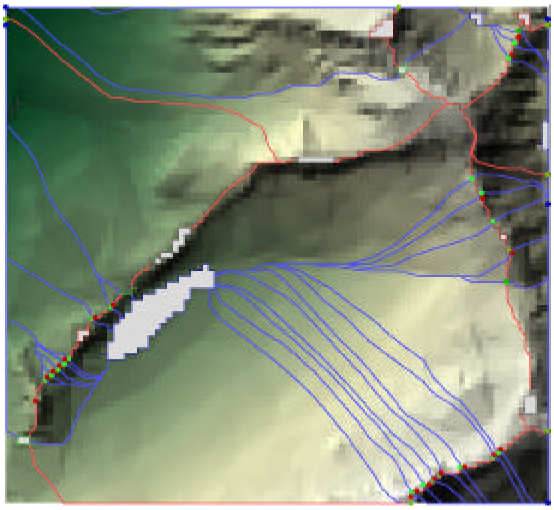

Figure 14. Surface network derived from the test area DEM.

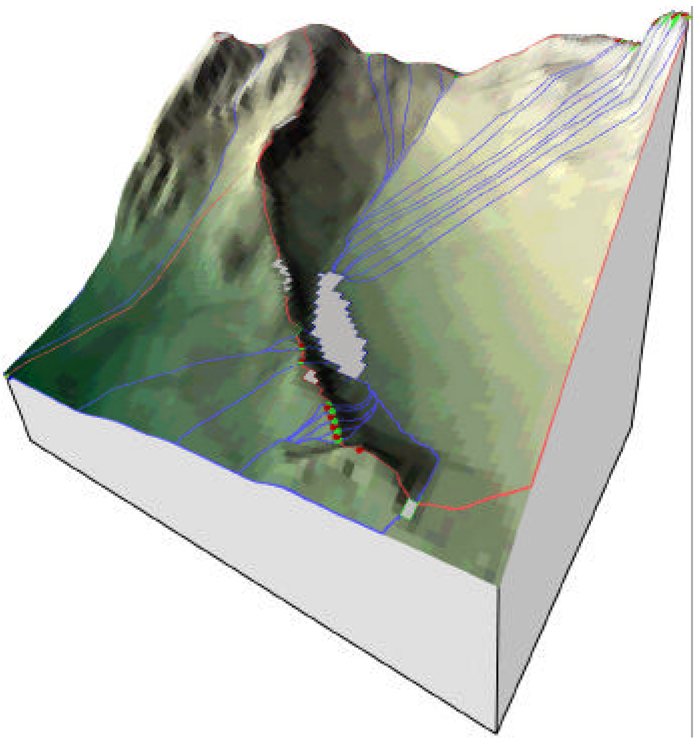

Figure 15. 3D view of the test area DEM and the according surface network.

A small number of horizontal areas are identified, some of them representing critical points. They are correctly incorporated in the surface network. Again, the extracted network is consistent and complete, no errors could be detected by visual inspection. However, as in the previous example, chains of peaks and passes occur along some ridges.

7. Conclusions and Outlook

The presented method for extracting surface networks from raster data has shown to produce consistent and complete results. The specification of the bilinear surface patches does not rely on user interaction and heuristics, and the network itself is mathematically well-defined. Consistency and completeness are achieved by considering the edge of the area of interest and by explicitly detecting and incorporating horizontal areas. Processing horizontal areas is the only step where applying a heuristic procedure is required. With respect to the various causes and interpretations of horizontal areas, this is considered a benefit rather than a drawback.

The method extracts all features of the real surface that are captured by the data. However, a potentially large number of spurious features are introduced by the bilinear interpolation scheme. Depending on the roughness of the terrain, the data sampling method, and the transformations applied to the data (e.g., smoothing), the raster data model may cause a considerable number of spurious pits and peaks. The bilinear interpolation scheme amplifies this effect. Additionally, the bilinear interpolation introduces passes along crests and valley lines. As a result, the surface network is often characterised by chains of passes and peaks along crests and by chains of passes and pits along valley lines.

There are two approaches to tackle this problem:

- Alternative interpolation schemes (e.g., cubic interpolation) reduce the generation of spurious passes. The methods for identifying critical points and tracing critical lines need to be adapted to the new interpolation scheme. For instance, pits and peaks may be located at raster points *and* inside cells.
- Creating hierarchies of drainage areas and mountains enables measuring the significance of each pass, which in turn allows identifying spurious passes. These passes may then be removed from the network with the help of homomorphic contractions (Pfaltz 1976, Wolf 1984, Rana and Wood 2000). This approach is subject of ongoing research by the author.

8. References

DE FLORIANI, L., and PUPPO, E., 1992, A hierarchical triangle-based model for terrain description. In Frank, A.U., Campari, I., and Formentini, U. (eds.), *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. Lecture Notes in Computer Science (Berlin-Heidelberg: Springer), pp. 236-251.

FLORINSKY, I.V., 2002, Errors of signal processing in digital terrain modelling. *International Journal of Geographical Information Science*, **16** (5), 475-501.

FOWLER, R.J., and LITTLE, J.J., 1979, Automatic extraction of irregular network digital terrain models. *Computer Graphics*, **13**, 199-207.

PEUCKER, T.K., and DOUGLAS, D.D., 1975, Detection of surface-specific points by local parallel processing of discrete terrain elevation data. *Computer Graphics and Image Processing*, **4**, 375-387.

PFALTZ, J., 1976, Surface Networks. *Geographical Analysis*, **8**, 77-93.
RANA, S., and WOOD, J.D., 2000, Weighted and metric surface networks – new insights and an interactive application for their generalisation in TCL/TK. *CASA Working Paper Series* (London: University College London), http://www.casa.ucl.ac.uk/surfacenetworks2.pdf
SCHNEIDER, B., 2001, On the Uncertainty of local shape of lines and surfaces. *Cartography and Geographic Information Science*, **28** (4), 237-247.
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. In Post, F. and Gobel, M. (eds.): *Eurographics '95*, Vol. 14 (Blackwell Publishers), pp. C181-C192.
WARNTZ, W., 1966, The topology of a socioeconomic terrain and spatial flows. *Papers of the Regional Science Association*, **17**, 47-61.
WERNER, C., 1994, Explorations into the formal structure of drainage basins. *Earth Surface Processes and Landforms*, **19**, 747-762.
WOLF, G.W., 1984, A mathematical model of cartographic generalization. In *Proceedings Geo-Processing 2* (Amsterdam: Elsevier Science Publishers), 271-286.
WOLF, G.W., 1990, Metric Surface Networks. In *Proceedings of the 4th International Symposium on Spatial Data Handling 1990*, Zurich, Switzerland, Vol. 2, pp. 844-856.
WOLF, G.W., 1991, Characterization of functions representing topographic surfaces. In *Proceedings Autocarto 10*, Baltimore, 186-204.
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.D., 2000, Constructing Weighted Surface Networks for the Representation and Analysis of Surface Topology. In *Proceedings 5th International Conference on GeoComputation*, Chatham, UK, http://www.soi.city.ac.uk/~jwo/GeoComputation00/.