Surface Topography: Metrology and Properties

Surfaces—topography and topology

Gert W Wolf
Department of Geography, University of Klagenfurt, Klagenfurt, Austria
E-mail: gert.wolf@aau.at

Keywords: topological data structures, surface characterisation, relevance of a topological feature, surface simplification, degree of simplicity, ISO 25178-2, ISO 16610-85

Abstract
ISO 25178-2 specifies areal field parameters as well as areal feature parameters. While for the first group the whole set of points defining a scale-limited surface is considered, for the second group only a subset of surface features is taken into account. As a consequence, an adequate data structure for surface characterisation in combination with an appropriate method for surface simplification are required. Three data structures for surface characterisation, namely, Morse-Smale complexes, weighted surface networks and change trees are discussed. Hereafter, the focus is laid on approaches for determining the relevance of topological features with respect to surface topography. Another topic of interest is surface simplification, i.e. the process of deriving from an original surface a second surface of decreased complexity, but with its structural properties being retained. Within the geosciences this concept is associated with the transition from large-scale maps to small-scale maps, whereas in the technical sciences it corresponds to the reduction of measurement noise. From a topological point of view, a theorem proven by Matsumoto may be regarded as the core theorem with respect to surface simplification. Its combination with the two concepts of relevance of topological feature and degree of simplicity represents the basis of a formal procedure for surface simplification as required in ISO 25178-2 and ISO 16610-85.

1. Introduction
Surfaces represent an important tool for the description and analysis of a multitude of phenomena related to the social, economic, natural and technical sciences. Evidently, this widespread use of surfaces has caused the need for a proper framework with regard to their adequate formal characterisation and simplification. As a consequence, researchers engaged in different disciplines, such as the geosciences, computer graphics or mathematics, have focused their interest on the development of new methods that are applicable to all types of surfaces. After a brief survey of the required mathematical concepts (section 2), three fundamental topics will be discussed, namely, the digital representation of surfaces (section 3), the relevance of topological features (section 4) and the simplification of surfaces in a digital environment (section 5). In section 6, finally, the relevance of the obtained results for the two standards ISO 25178-2 [1] and ISO 16610-85 [2] will be exposed.

2. Mathematical background
At the beginning of this section, it should be emphasised that surfaces do not exist per se, but they are merely models. According to Stachowiak [3], models can be characterised in a threefold way, namely, by a mapping feature (a model is always a model of something, it is a representation of a natural or an artificial original), by a reduction feature (a model does not capture all the attributes of the original, but only those that appear relevant to its creator) and by a pragmatic feature (a model is not clearly assigned to its original, but it is constructed for certain subjects, for certain time intervals and it is restricted to certain mental or actual operations). With this in mind, the formal characterisation of a surface will be given next.

A surface is generally considered as a function $f(x, y)$ that assigns to each point $(x, y)$ of a domain either its altitude in case of a topographic surface or an arbitrary data value otherwise. With respect to the theory presented in the following, two assumptions
regarding \( f(x, y) \) have additionally to be made, viz., (a) that \( f(x, y) \) is a smooth function, i.e. it is infinitely differentiable, and (b) that \( f(x, y) \) has only nondegenerate critical points, namely, the well-known local minima (pits), saddle points (saddles, passes) and local maxima (peaks), as portrayed in figure 1 (figure 1 was created with Wolfram Mathematica® 10).

Without going into details, it should be noted that a critical point \((x_0, y_0)\), which is also termed surface-specific point in this context, can be characterised by the first-order and second-order partial derivatives of \( f(x, y) \) at this point (an introduction into multivariable calculus can be found, for instance, in Ghorpade and Limaye [4] or Moskowitz and Paliogiannis [5]).

Functions satisfying the two previously mentioned conditions, namely, to be smooth and to have only nondegenerate critical points, are termed Morse functions (for an introduction into Morse theory confer, for example, Matsumoto [6], Milnor [7] or Zomorodian [8]). As the approaches presented within this paper are based on Morse functions, their most relevant features in combination with two important theorems will be specified next.

The first property worth mentioning is that the critical points of Morse functions are always isolated. The second property is, that in case they are defined on a compact manifold, they have only finitely many critical points. The third property is that any smooth function can be closely approximated by a Morse function [9, 10], whereby this feature makes Morse functions particularly suitable for analysing real-world data.

The first theorem, the Morse Lemma, states that in the neighbourhood of a local minimum a two-dimensional Morse function always looks like an upright paraboloid, in the neighbourhood of a saddle point always like a hyperbolic paraboloid and in the neighbourhood of a local maximum always like an inverted paraboloid. According to the Morse Lemma, the critical points of a Morse function can thus only look like the ones illustrated in figure 1.

The second theorem, the Morse inequalities, states, amongst others, that for all Morse functions that are defined on the two-dimensional surface of a sphere the relationship number of pits minus number of passes plus number of peaks equals two holds. As Maxwell [11] has shown, this equality remains valid if the Morse functions are defined on a simply connected domain that is bounded by a closed contour line.

3. Digital representation of surfaces

Topographic data and in many cases scientific data, too, consist of measurements taken over a geometric domain, thus representing a discrete sample of points obtained from a continuous function, which is defined on a two-dimensional surface, such as the Earth’s surface or a plane. As these functions are in general too complex to be expressed in a closed form, GIScientists, mathematicians and computer scientists have developed a multitude of data structures for their adequate representation. At present, different classification schemes exist, according to which the designed data structures can be grouped in dependence of some predefined criteria. One major distinction in this context is based on the type of data points being applied within the data structure, either arbitrary data points or the surface-specific points only. A survey of the most relevant spatial data structures according to this classification scheme can be found in figure 2, with the numbers referring to the sections in which they are discussed. Regardless of the chosen data structure, a common characteristic of all of them is that they take both topological and geometrical features into account. While the topological properties are used for describing the relationships between the data points, the geometrical ones are required for embedding the topological structures into a metric space.

3.1. Digital representation of surfaces by arbitrary data points

When applying arbitrary data points for surface characterisation, terrain is described by the altitudes of these points, which are arranged in the \((x, y)\)-plane either in the form of a regular grid or a Triangular Irregular Network (TIN), both of which are shown in figure 3 (figure 3 was created by means of Wolfram Mathematica® 10).
In a regular grid the data points are organised as a rectangular array (figure 3(a)) that has to be fine enough to ensure an appropriate approximation of the surface everywhere. In the following, the altitudes of surface points being different from the grid points can be approximated by various methods, such as Bézier surfaces, B-spline surfaces or bicubic Hermite patches \[12, 13\]. Considering the pros and cons of regular grids, it can be said that their main advantage is their easy handling, which is due to the grids’ inherent topology providing fast access to information related to neighbouring data points via a simple indexing technique. Conversely, the biggest disadvantage of regular grids results from the fact that the roughness of terrain changes from one landform to the other. For a pre-specified level of precision, the grid must therefore be adjusted to the roughest terrain (generally mountainous regions), with the effect that the grid may be highly redundant in areas where terrain is less rough or even smooth (usually flat plains) \[14, 15\].

In contrast, Triangulated Irregular Networks do not offer any kind of regularity, but approximate terrain by triangular patches, whereby \textit{a priori} no restriction with respect to the selected data points is made. Evidently, the projection of the vertices and edges of the triangle patches into the \((x, y)\)-plane results in a planar graph, defining a tessellation of the plane into triangular patches, called \textit{two-dimensional triangulation} (figure 3(b)) \[14, 15\]. Depending on the chosen type of triangulation, different methods for approximating the altitudes of those surface points, which are different from the triangle vertices, do exist. In the case of a regular triangulation, Bernstein polynomials, B-splines or Coons-Gordon surfaces may be applied \[12, 13\], whereas in the case of an irregular triangulation over each triangle a function satisfying certain pre-specified conditions, such as the Clough-Tocher scheme or the Powell-Sabin scheme, is constructed \[12\].

For the sake of completeness, it should be pointed out that more recent research has focused its interest...
on the development of hybrid approaches, which combine the regular grid structure with Triangulated Irregular Networks [16, 17].

3.2. Digital representation of surfaces by surface-specific points

Another option for surface characterisation is by means of their critical points. Although well established in several disciplines, such as computational geometry or computer graphics, these concepts can in general be considered as exceptions and not as a rule. Within the geosciences Peucker [18] was the first to adopt this approach, thereby emphasising that the information content of a surface-specific point is significantly higher than that of any other surface point, as this type of point provides information not only about a particular location but also about its neighbourhood; within surface metrology Scott [19] may be considered as the pioneer with respect to the application of critical points for surface texture characterisation.

In the remainder of this section, three approaches for the topological characterisation of a surface by their critical points will be presented, namely, Morse-Smale complexes, weighted surface networks and contour trees (change trees). Thereby it will be assumed that the critical points and critical lines (surface-specific lines), i.e. the lines connecting the critical points, are known. For their identification and for the identification of other surface features several algorithms were developed, which may be classified according to the specific tasks they are designed for in the following way: (a) extraction of the critical points from a given set of data points, which are usually arranged in form of a rectangular grid [20–22]; (b) detection of the critical lines, i.e. the ridges and valleys, whereby the algorithms are either based on the given data points only [23–28] or they are organised in such a form that they interpolate, first of all, the surface by a function of two variables and identify hereafter the ridges and valleys [29–32]; (c) identification of regions of interest [33, 34]; (d) extraction not only of the critical points and critical lines, but of complete topological structures, such as surface networks [35–39] or contour trees [38–50].

3.2.1. Morse-Smale complexes

The terms Morse complex and Morse-Smale complex (the definitions are not unique, but depend on the particular author) can be traced back to Maxwell [11], who proposed the subdivision of a landscape into regions that are either hills or, alternatively, dales (basins), with each point of the landscape belonging to a specific hill as well as to a specific dale. More than a century later Maxwell’s approach was taken up by various scientists [8, 51–65], who formalised his concepts and applied them primarily within computational geometry and computer graphics.

The basic concept underlying Morse-Smale complexes is that of an integral line, which is, broadly speaking, the line of steepest ascent between two points, termed origin and destination, on a smooth compact n-dimensional manifold. Due to its definition, the following four properties of integral lines always hold: (a) two integral lines either have disjoint images or they are the same; (b) apart from the critical points, the entire manifold M is covered by the images of the integral lines; (c) both the origin and the destination of an integral line are always critical points; (d) critical points are images of constant integral lines by themselves [8, 51–53, 59–61]. Figure 4 visualises the graph of the function \( f(x, y) = 2 \sin x + 2 \sin y \) for \( x, y \in [-3\pi, 3\pi] \) together with some integral lines.

Between contour lines and integral lines the following relationships exist: (a) while contour lines represent curves of equal altitude, integral lines represent curves that follow the direction of steepest ascent; (b) integral lines and contour lines are perpendicular to each other; (c) for all non-critical points \( p \) of a smooth compact manifold are the contour line and the integral line passing through \( p \) well defined [57].

Integral lines represent the basic concept for the formal definition of ascending manifolds, also termed unstable manifolds, and descending manifolds, also termed stable manifolds. Precisely speaking, an ascending manifold \( A(p) \) of a critical point \( p \) is the set of all points belonging to integral lines whose origin is \( p \), whereas a descending manifold \( D(p) \) of a critical point \( p \) is the set of all points belonging to integral lines whose destination is \( p \). If a Morse function is defined on a two-dimensional manifold, its ascending manifolds are obviously equivalent to the basins of the corresponding surface, while the descending manifolds are equivalent to its hills. For theoretical reasons, the boundary of the surface is represented either by a virtual pit of altitude \( -\infty \) or a virtual peak of altitude.
+, a process that is termed one-point compactification [8, 61, 63].

Figure 5 visualises the graph of the function \( f(x, y) = 2 \sin x + 2 \sin y \) for \( x, y \in [-3\pi, 3\pi] \) in combination with its subdivision into (a) ascending (unstable) manifolds and (b) descending (stable) manifolds.

Figure 6. (a) Two manifolds \( A \) and \( B \) that intersect transversally and (b) two manifolds \( A \) and \( B \) that do not intersect transversally.

Figure 7. Graph of the function \( f(x, y) = 2 \sin x + 2 \sin y \) for \( x, y \in [-3\pi, 3\pi] \) together with its Morse-Smale complex.

Morse-Smale functions are Morse functions, whose ascending and descending manifolds intersect only transversally, i.e. the tangent vectors in the points of intersection are defined and not multiples of each other. Less formally, transversality refers to the fact whether two manifolds are `parallel' at their intersection point or not. In the two-dimensional case transversality entails that the ascending and descending one-dimensional manifolds cross, with their point of intersection always being a saddle [8, 51, 60–62].

For a Morse-Smale function \( f \), the connected components of \( A(p) \cap D(q) \) for all critical points \( p, q \) of a manifold are referred to as Morse cells. The cells of dimension zero, one and two are called vertices, arcs and regions, their collection is termed as the Morse-Smale complex of \( f \).

Figure 7 displays the graph of the function \( f(x, y) = 2 \sin x + 2 \sin y \) for \( x, y \in [-3\pi, 3\pi] \) in combination with the corresponding Morse-Smale complex; obviously, each vertex represents a critical point, each arc is half of an ascending or a descending one-dimensional manifold of a saddle and each region
is a component of the intersection of an ascending two-dimensional manifold of a minimum and a descending two-dimensional manifold of a maximum (figure 7 was created with Wolfram Mathematica® 10).

For Morse-Smale complexes the Quadrangle Lemma holds, which states that all regions of a Morse-Smale complex are quadrangles, with each quadrangle being bounded by a pit, a pass, a peak and a further pass in this succession (confer figure 7) [8, 60].

Another example of a Morse-Smale complex is given in figure 8, illustrating a surface containing pits $x_i$ ($i = 1, \ldots, 4$), passes $y_j$ ($j = 1, \ldots, 11$) and peaks $z_k$ ($k = 1, \ldots, 9$) together with the connecting course lines and ridge lines (figure 8 was constructed with AutoCAD® Civil 3D® 2019). The heights of the data points are $h(x_1) = 70$, $h(x_2) = 30$, $h(x_3) = 20$, $h(x_4) = h(x_5) = h(x_6) = h(x_7) = h(x_8) = h(x_9) = h(x_{10}) = h(x_{11}) = h(x_{12}) = h(x_{13}) = h(x_{14}) = h(x_{15}) = h(x_{16}) = 0$, $h(y_1) = 460$, $h(y_2) = 450$, $h(y_3) = 350$, $h(y_4) = 130$, $h(y_5) = 120$, $h(y_6) = 100$, $h(y_7) = 90$, $h(y_8) = 150$, $h(y_9) = 80$, $h(y_{10}) = 60$, $h(y_{11}) = 40$, $h(z_1) = 850$, $h(z_2) = 650$, $h(z_3) = 480$, $h(z_4) = 450$, $h(z_5) = 380$, $h(z_6) = 370$, $h(z_7) = 220$, $h(z_8) = 220$ and $h(z_9) = 220$ (for a better readability the data points $x_1, \ldots, x_{16}, y_1, \ldots, y_{11}, z_1, \ldots, z_9$ are denoted by $x_1, \ldots, x_{416}, y_1, \ldots, y_{11}, z_1, \ldots, z_9$ throughout the figures of this paper).
Figure 8(a) visualises the topography of the surface by spline-interpolated contour lines, while figure 8(b) depicts its critical points in combination with the connecting courses (marked in blue) and ridges (marked in red). Evidently, the red lines mark the ascending cells of the saddle points, while the blue lines mark their descending cells, with the subdivision thus obtained forming the Morse-Smale complex of the surface. It should be noted that, according to the geometrical-topological nature of the approach, the virtual pit $x_0$ is counted as a single point in the topological data structure (Morse-Smale complex), while sixteen different data points $x_{4i}, x_{4i+1},...,x_{4n}$ are used in the associated geometrical data structure (embedding of the Morse-Smale complex in a coordinate system).

When ignoring the regions of the Morse-Smale complex and considering merely its vertices and arcs, the structure thus obtained is called the critical net (1-skeleton) of the Morse-Smale complex $[51, 56, 57]$. According to Biasotti et al $[51]$, the graph representation of the critical net is equivalent to the corresponding weighted surface network, a topological data structure that will be described in the next section.

It is worth mentioning that the subdivision of surfaces either into ascending or descending manifolds or into quadrangular regions was also suggested by scientists conducting research in other disciplines—however, without employing the terminology introduced in this section. In surface metrology, for instance, the segmentation of a surface into mutually exclusive hills or basins, which are called areal motifs, was first described by Scott $[19, 66-68]$, whose concepts have found their way into the two standards ISO 25178-2 $[1]$ and ISO 16610-85 $[2]$.

### 3.2.2. Weighted surface networks
Maxwell’s approach $[11]$ was put into a graph-theoretic framework by Pfaltz $[69, 70]$, who gave the formal definition of a surface network, a data structure for the topological characterisation of surfaces. The vertex set of this type of graphs consists of the critical points, the edge set of the critical lines of a Morse function that is defined over a domain, which is simply connected and bounded by a closed contour line. A weighted surface network represents an extension of a surface network in such a way that, in addition, positive real numbers are assigned to the edges and nodes to denote their relevance with respect to the topography of the corresponding surface.

Figure 9 visualises the weighted surface network belonging to the surface depicted in figure 8. Evidently, the vertex set is composed of the three subsets $P_0$ (the set of all pits), $P_1$ (the set of all passes) and $P_2$ (the set of all peaks), while the elements of the edge set belong either to the subgraph $[P_0, P_1]$ (representing the course lines) or to the subgraph $[P_0, P_2]$ (representing the ridge lines). The edge weights specify height differences between adjacent critical points and are defined as $h(y_i) - h(x_i)$ and $h(z_i) - h(y_i)$ respectively, whereby $h$ indicates the altitude of a critical point (vertex weights will be defined later).

In order to be regarded as an abstract representation of the topological structure of a surface, a weighted surface network, also termed Pfaltz graph, has to satisfy the properties listed below (for the formal definition confer Pfaltz $[69, 70]$ or Wolf $[71-76]$; an introduction into graph theory can be found, for instance, in Bondy and Murty $[77]$): (a) It must be a planar graph, because an intersection of its edges, being equivalent to an intersection of the corresponding ridges and courses of the surface, would imply the impossibility of the realisation of the latter; note that the graph depicted in figure 9 is a planar graph, as it can also be drawn in such a way that its edges do not cross. (b) The two subgraphs $[P_0, P_1]$ and $[P_0, P_2]$ must be connected, thus reflecting the fact that the corresponding pits and saddles are connected by course lines and the corresponding saddles and peaks are connected by ridge lines. (c) The number of pits minus the number of passes plus the number of peaks must always be two, a result that is due to the Morse inequalities. (d) The indegree of each pass must be equal to its outdegree, namely two, thus ensuring that all saddle points are nondegenerate. (e) The prerequisite that the existence of a path from pit $x$ via pass $y$ to peak $z$, whose edges have only valency one, induces the existence of another path from pit $x$ to peak $z$ via a distinct saddle $y$, whose edges also have only valency one, is derived from the fact that the surface under consideration is a Morse-Smale complex to which the Quadrangle Lemma applies. (f) The assumption that $(x, y)$ is an edge of a circuit in the subgraph $[P_0, P_1]$ if and only if $\text{val}(y, z) = 2$ for all peaks excludes edge configurations corresponding to non-realisable configurations of the respective surface-specific lines; evidently, an analogous assumption must hold for the edges of the subgraph $[P_0, P_2]$. (g) The edge-weights must be positive real numbers and have thus to be specified as $h(y_i) - h(x_i)$ and $h(z_i) - h(y_i)$ respectively. (h) All paths from a pit $x$ to a peak $z$ must exhibit the same height difference, no matter which saddle point is passed. (i) If two edges are leading from a pit $x$ to a pass $y$, then to both of them the same edge weight must be assigned, because all course lines between a pit and a pass exhibit the same height difference; evidently, an analogous assumption must be valid for edges from a pass $y$ to a peak $z$.

For the sake of completeness, it should be pointed out that, according to Pfaltz $[69, 70]$, a (weighted) surface network is a directed graph, with each edge pointing from the lower located vertex to the higher located one, thus inducing that the edges may be regarded as abstract representations of the corresponding integral lines. From a graph-theoretic point of view, however, this presumption is not necessary, because the edge directions can be inverted or even omitted.

It is worth mentioning that despite their different appearance the two graphs depicted in figures 8(b) and 9...
are topologically equivalent since their vertex sets and edge sets are identical. As can be seen from figure 9, in the case of surface networks, too, a virtual pit, in this instance also termed surrounding pit, is used as a one-point compactification, thereby representing the boundary of the surface.

Although the explanations in section 3.2.2 are written in such a way that the features of a weighted surface network are derived from the properties of a Morse-Smale complex (confer section 3.2.1), it should be emphasised that weighted surface networks were developed as a data structure of its own [69–73]. Practical applications of surface networks can be found primarily within GIScience [21, 22, 35–37, 71, 72, 78–85].

3.2.3. Contour trees (change trees)

It can be assumed that Boyell and Rushton [86] were the first to apply a tree-like data structure, which they called enclosure tree, for representing containment relationships among the contours of topographic maps. Today’s preferred term contour tree was introduced by Boehm [87], who compared different storage methods related to surface topography. In the late 1970s, Mark [88, 89] applied the surface tree for the analysis of geomorphic surfaces and about a decennium afterwards, Kweon and Kanade [90, 91] introduced the (topographic) change tree into pattern recognition and image analysis.

In the years that followed numerous papers were published which focussed, on the one hand on the design of efficient algorithms for the extraction of contour trees [38–50] and on the other hand on the conception of special types of contour trees like the augmented contour tree [40, 92], the multiresolution contour tree [46], the subdomain-aware contour tree [93], the branch decomposition tree [46, 50, 94, 95] or joint contour nets [64].

At present, computational geometry and computer graphics can be regarded as the main application areas of contour trees, with their application range comprising the identification of surface-specific points, the analysis of volume data, the extraction of isosurfaces as well as the visualisation of complex phenomena [43, 51, 64, 96].

Although contour trees constitute a profoundly geographic data structure, they are rather infrequently used within the geosciences. Up to now, only a couple of authors have employed contour trees—for instance, for the geomorphic analysis of surfaces [88, 89], the analysis of data fraught with uncertainty [95], the automatic labelling of contour lines [97, 98], the description and storage of terrain at different levels of detail [99, 100] or the detection and analysis of depressions [101].

The change tree has also attained a lot of interest in surface metrology due to the approach of Scott [19, 66–68], who applied it for a syntactical relational description of surface texture. Scott’s central assumption in this context concerns the human comprehension of surface topography, stating that this comprehension is not based on numerical data, but on patterns of features, such as those of the surface-specific points and surface-specific lines together with the interlinking relations. Formally, Scott took up the
concept of profile motifs and transferred it to areal motifs by applying Maxwell’s ideas [11], but redefining them in terms of modern topology and calculus.

Figure 10 shows two visualisations of the contour tree corresponding to the surface depicted in figure 8. Figure 10(a) portrays the contour tree in such a way that the elevations of the surface-specific points can be read from the vertical line on the left, while in figure 10(b) the height differences between two adjacent critical points are specified by edge-weights. For the sake of completeness, it should be pointed out that only very few authors prefer vertex-weights indicating absolute elevations to edge-weights denoting relative ones [43, 95].

As can be seen from figure 10, a contour tree, which visualises a surface by capturing its topological evolution when the isovalue varies, satisfies the following properties: (a) The vertices represent those contour lines that pass through the critical points of the Morse function \( f \) characterising terrain; to be more precise, vertices of degree one represent contour lines containing local minima or local maxima (in these points, contours are created or destroyed), whereas vertices of degree three or higher represent contour lines containing saddle points (in these points, two or more contours merge or one contour splits into two or more components). If all saddles have, in addition, distinct function values, their degrees will always be three [48], as in the tree depicted in figure 10. (b) The edges represent those contours that do not contain a critical point of \( f \), whereby it is assumed that each edge is directed from the vertex with the higher elevation to the vertex with the lower elevation (in this context it is noteworthy that the vertices and edges of the contour tree are also referred to as supernodes and superarcs in order to make a distinction between them and the infinite number of nodes that lie on the edges of the contour tree and represent contours containing no critical point [31]). (c) As in the case of Morse-Smale complexes and weighted surface networks, the points that are located outside the domain of the Morse function defining the surface are represented by a virtual pit [38, 39, 67, 102], thereby assuming that there exists no other critical point outside the domain of the Morse function and the virtual pit’s elevation is \(-\infty\) (in figure 10 the dashed arrow being directed to the virtual pit \( x_4 \) indicates that the respective edge is of infinite length). A dual analogon to the virtual pit also exists, viz., that of the virtual peak, with the only difference being the assumption that the virtual peak’s elevation is \(+\infty\). For Scott [67] the concepts of the virtual pit and virtual peak constitute an application of Ockham’s Razor, a philosophical assertion maintaining plainness in theory design by stating non sunt multiplicanda entia praeter necessitatem, i.e. entities are not to be multiplied beyond necessity.

In conclusion, it should be mentioned that, according to several authors [59, 61, 63], a major drawback of contour trees—as compared to Morse-Smale complexes or weighted surface networks—is that they contain no geometrical information related to the lines of steepest ascent and steepest descent, respectively.

**Figure 10.** Two different visualisations of the contour tree associated with the surface portrayed in figure 8(a); the dashed arrow pointing to \( x_4 \) indicates that this vertex represents the virtual pit and that the respective edge has infinite length. (a) A combined geometrical-topological representation (the geometrical aspect refers to the vertical axis on the left indicating altitude) and (b) a purely topological visualisation (edge-weights specify differences in altitudes between two adjacent critical points).

4. The relevance of topological features

After the discussion of various data structures for surface characterisation, the present section is devoted to the topic of how the relevance of the critical points
and critical lines with regard to topographic terrain can be measured. Although all existing concepts are based on height differences between surface-specific points, the approaches vary slightly in dependence of the chosen data structure.

4.1. Persistence

In Morse-Smale complexes the numerical measure specifying the relevance of topological features is termed (topological) persistence [8, 51–53, 58, 60–65]. It relies on the fact that the behaviour of contours changes merely in the critical points; more precisely, in a pit a new contour line is created, in a pass contour lines merge or split, while in a peak a contour line is destroyed. With respect to surface analysis, the lifetime between birth and death of a topological feature is deemed as meaningful, whereby features having a long lifetime are considered to be important, while features having a short lifetime are considered to be unimportant. Lifetime itself is measured in terms of height differences between the two critical points that create and destroy the respective feature, with these two points being referred to as persistence pair. (For the sake of completeness, it should be pointed out that in literature persistence is sometimes assigned to the topological feature itself, sometimes to the critical points characterising the topological feature and sometimes to one of the critical points of the persistence pair.)

Evidently, in the case of a two-dimensional surface each persistence pair consists either of a pit and a pass or of a pass and a peak, while the persistence is represented by the height difference between these two points. When determining the persistence of the topological features of a surface, two observations are noteworthy, namely, (a) the persistence is not computed for all pairs of adjacent critical points in a Morse-Smale complex (adjacent critical points are those that are connected either by a course line or a ridge line), but only for a small subset of them, and (b) two critical points forming a persistence pair must not necessarily be connected by a course line or a ridge line in the Morse-Smale complex [60, 65].

Figure 11 visualises the persistence of the topological features of the surface depicted in figure 8 by barcodes, which is a very common and very illustrative method for this purpose.

In the meantime the concept of persistence was modified in several ways, thereby taking a lot of additional features, such as the following ones, into account:

(a) The scale space persistence is based on the deep structure of the scale space, which is a one-parameter family of functions, obtained by cumulative smoothing of the initial function. It allows
the tracing of the critical points through different scales and takes, in addition, the spatial extent of a critical point into account [103].

(b) The separatrix persistence facilitates the differentiation between salient and non-salient separatrices, i.e. between important and unimportant courses and ridges [65, 104]. This is achieved by assigning a weight to each point along a separatrix, thus taking into account that the importance of a feature may change smoothly along the line.

(c) The topological saliency does not only consider the size of the topological features, but, in addition, their distribution in space [58]. Accordingly, the topological saliency may be considered as the relative importance of a topological feature with respect to other features within a specified neighbourhood.

4.2. The importance of surface-specific points in weighted surface networks

In weighted surface networks, two types of numerical values exist, viz., edge weights and vertex weights, to denote the importance of topological features with respect to surface topography. Edge weights \(w(e_i)\), which are assigned to all course lines and ridge lines, specify height differences between the two critical points being connected by the respective line. Vertex weights \(I(v_i)\), which are assigned to all pits and peaks, resemble persistence in this respect that critical points with a high \(I\)-value are seen as important, whereas critical points with a low \(I\)-value are considered as unimportant. Vertex weights themselves can be derived from the edge weights in several ways, such as the ones listed below, thereby facilitating the consideration of the specific character of the problem under discussion [71,75,76]:

(a) The minimum of all height differences between a pit (peak) and its adjacent passes represents the most commonly used indicator for the relevance of a surface-specific point with regard to topography.

(b) The maximum of all height differences between a pit (peak) and its adjacent saddles constitutes an alternative numerical measure.

(c) The sum of the height differences of all courses (ridges) being incident with a pit (peak) induces that those critical points with a large number of adjacent passes, thus representing critical points with a high level of centrality, will, in general, be considered as the important ones.

(d) Propositions (a)–(c) can be modified by also including the absolute elevations of critical points, thus obtaining vertex weights that reflect the fact that the same height difference is of varying significance in mountainous regions and in plains.

When applying the first of the previously mentioned criteria for determining the importance of the critical points of the surface depicted in figure 8, one obtains the arrangement \(I(x_i) = 10, I(z_j) = 20, I(z_k) = 20, I(x_l) = 30, I(z_m) = 30, I(x_n) = 40, I(x_o) = 70, I(z_p) = 90, I(z_q) = 90, I(z_r) = 180, I(z_s) = 190, I(z_t) = 330\) and \(I(z_u) = 400\).

4.3. The importance of surface-specific points in contour trees (change trees)

In contour trees, course lines and ridge lines correspond to uniquely defined paths [38, 39] that can, however, not easily be identified. As a consequence, the calculation of their height differences as well as the computation of numerical values serving as measure for the relevance of the pits and peaks is much more complicated than in surface networks [76]. Strictly speaking, only one measure exists that can be calculated without great effort, viz., the minimum of all height differences between a pit (peak) and its adjacent passes. This property is due to the fact that the courses (ridges) leading from a pit (peak) to its lowest (highest) adjacent pass are equivalent to those edges in a contour tree that lead from the vertices of degree one, representing the pits and peaks, to the vertices of degree three (or higher), representing the passes. In the following, the obtained values can be assigned either to the pits and peaks as vertex weights or to the edges leading from the pits and peaks to the respective passes as edge weights. Although this inconsistency, which is comparable with the inconsistency when assigning the persistence to a topological feature, may be considered as unsatisfactory, it is of no influence to practical applications.

When starting from figure 10(b) and determining the importance of the critical points of the surface depicted in figure 8, one obtains, in dependence of whether the importance is assigned to the vertices or to the edges, the arrangement \(I(x_i) = I(y_j, x_k) = 10, I(z_j) = I((z_m, y_l)) = 20, I(z_k) = I((z_m, y_l)) = 20, I(x_l) = I((y_j, x_k)) = 30, I(z_m) = I((z_m, y_l)) = 30, I(z_n) = I((y_l, z_k)) = 70, I(z_j) = I((z_m, y_l)) = 90, I(z_k) = I((z_m, y_l)) = 90, I(z_m) = I((z_m, y_l)) = 180, I(z_n) = I((z_m, y_l)) = 190, I(z_k) = I((z_m, y_l)) = 330, I(x_l) = I((z_m, y_l)) = 400\) and \(I(x_k) = I((y_j, x_k)) = −\infty\). It should be noted that, with the exception of the virtual pit, this arrangement coincides with the \(I\)-values listed in section 4.2.

5. The simplification of surfaces

Surface simplification describes the process of deriving from an original surface a second surface of decreased complexity, but with its structural properties being retained. Within the geosciences this concept is
associated with map generalisation, the transition from large-scale to medium-scale or small-scale maps, while in the technical sciences it is related to the reduction of measurement noise.

Depending on which data structure has been chosen for the digital representation of a surface, different simplification methods are available at present.

In case that the surface is represented by arbitrary data points, all available algorithms are geometrically driven, which implies that the extent of the simplification process is exclusively supervised by the magnitude of a numerical value. The mathematical bases of these approaches, which are the most commonly used today, are, for instance, the two-dimensional fast Fourier transform \([105]\), two-dimensional biorthogonal average interpolating wavelets \([106]\) or B-splines \([107]\).

In contrast, if the topography is characterised by the surface-specific points, a geometrical-topological approach is required, with the topological features supervising the extent of the generalisation process and the geometrical properties being responsible for a proper embedding of the topological structures into a metric space. While geometrical-topological approaches are well established, for instance, in computational geometry or computer graphics, they are very rarely used within the geosciences \([18, 69–71, 73–75, 108, 109]\) or within surface metrology \([66–68]\).

5.1. The Theorem of Matsumoto

The Theorem of Matsumoto, which also applies in the \(n\)-dimensional case \([6]\), may be regarded as the core theorem with respect to the simplification of the topological structure of surfaces that are characterised by their critical points. Applied to two-dimensional surfaces, it says, in clear and simple terms, that from a topological point of view the only valid simplification is the pairwise elimination of a pit together with its lowest adjacent pass or of a peak together with its highest adjacent pass.

The importance of Matsumoto’s Theorem results from its universality and the resulting independence from a chosen data structure for surface representation. Thus it can be considered as a generalisation of several theorems that were proved over the past decades for specific contractions being defined for Morse-Smale complexes, weighted surface networks and contour trees, thereby comprising the following approaches:

(a) Cancellations being defined for Morse-Smale complexes \([8, 51–55, 58, 60–62, 64, 65, 104]\).

(b) \(w\)-contractions being defined for weighted surface networks \([71–75]\).

(c) Wolf pruning being defined for contour trees (change trees) \([40, 41, 66–68, 96]\).

A detailed review of these contractions can be found in Wolf \([76]\), a visualisation of the interrelationships between the data structures, the associated numerical measures for the relevance of the topological features and the corresponding contractions is given

![Figure 12. Visualisation of the interrelationships between the data structures, the associated numerical measures for the relevance of the topological features and the corresponding contractions, with the numbers indicating the sections in which the topics are discussed.](image-url)
in figure 12, with the numbers indicating the sections in which the topics are discussed.

Figure 13 visualises the surface of figure 8(a) after the elimination of the pit \( x_1 \) in combination with the pass \( y_9 \), by spline-interpolated contour lines. (figure 13 was created with AutoCAD® Civil 3D® 2019). In the present case, the pit \( x_1 \) and, according to the Theorem of Matsumoto, its lowest adjacent pass \( y_9 \) were removed.

Figure 14 illustrates the topological data structures associated with the simplified surface shown in figure 13. Figure 14(a) depicts the critical points of the Morse-Smale complex in combination with the connecting course lines (marked in blue) and ridge lines (marked in red), figure 14(b) visualises the contracted weighted surface network and figure 14(c) displays the pruned contour tree (change tree); concerning the last, it should be noted how the branch between \( x_1 \) and \( y_9 \) shown in figure 10 is simply cut off, i.e. pruned, thus resulting in the tree illustrated in figure 14(c).

5.2. An algorithm for surface simplification
Matsumoto’s Theorem ensures that every two-dimensional surface can be simplified by repeated elimination of either a pit together with its lowest adjacent pass or of a peak together with its highest adjacent pass. Provided that a surface has \( m \) saddle points, this elimination process will always terminate after \( m - 1 \) steps, when a surface consisting either of one pit, one pass and two peaks or of two pits, one pass and one peak is obtained.

Although of theoretical interest, an elimination process as previously described is of no practical use for real-world applications, which is why it has to be modified in at least two respects. First of all, the elimination of the pits and peaks must not occur in an arbitrary order, but according to their relevance (confer section 4), with the unimportant features being removed first and the important features being maintained. Secondly, as a surface containing only four critical points is oversimplified for any practical application, different criteria have to be specified to ensure the termination of the contraction process whenever desired, thereby obtaining a surface with a predefined degree of simplicity. Some possible termination rules are listed below [76]:

(a) The number of critical points (pits, peaks) that should be retained in the simplified surface.
(b) The proportion of critical points (pits, peaks) that should be retained in the simplified surface in relation to the number of critical points (pits, peaks) of the original surface.
(c) The height difference between a pit (peak) and its (lowest, highest) adjacent pass that should be retained in the simplified surface.
(d) The proportion of the height difference between a pit (peak) and its (lowest, highest) adjacent pass that should be retained in the simplified surface in relation to a predefined value being derived from the elevations of the original surface.

The combination of the two concepts of relevance of a topological feature and degree of simplicity with the
assertion of Matsumoto’s Theorem enables the specification of an algorithm for surface simplification. As can be seen from line 5 of algorithm 1, the surrounding pit (peak) must not be removed by the contraction process, because its elimination would entail the indefiniteness of the study area.

Algorithm 1. (Simplification of a surface)

1: procedure SIMPLIFICATION
2: 
3: Define the desired degree of simplicity for the contracted surface
4: 
5: while the specified degree of simplicity is not achieved do
6:     Compute the relevance of the pits $x$ and peaks $z$
7:     Identify the pit $x^o$ or the peak $z^o$, whose relevance is minimal and which is situated within the boundary contour
8:     Eliminate the pit $x^o$ or the peak $z^o$ together with its (lowest, highest) adjacent pass; update the differences in altitudes appropriately
9: end while
10: Stop
11: end procedure

Algorithm 1 may be regarded as generalisation of several procedures that were developed in the past decennia to contract topological data structures based on surface-specific points, such as Morse-Smale complexes, weighted surface networks or contour trees (for a detailed review confer Wolf [76]).

Figures 15 and 16 demonstrate the modus operandi of algorithm 1 when applied to a topographic surface and a scale-limited surface, respectively (for a detailed discussion of the latter confer section 6). Figure 15(a) visualises the original surface, located in the Latschur Mountains, Austria. A simplified version of this surface, which was obtained by eliminating twenty peaks together with their highest adjacent passes, is illustrated in figure 15(b); the contour lines are spline-interpolated (figure 15 was created with AutoCAD® Civil 3D® 2019).

Figure 16(a) gives a three-dimensional view of a scale-limited surface, while figure 16(b) visualises its motifs. Figure 16(c) sketches the simplified surface after having performed a pruning with a factor of 5% with respect to the height difference between the highest and lowest points of the surface, whereas figure 16(d) illustrates the simplified surface after having performed a pruning with a factor of 10%. As can be seen from figure 16, the number of motifs is reduced drastically by the simplification process, in the present case from 1308 in the original surface to 387 in the surface that was pruned with a factor of 5% and, finally, to 202 in the surface that was pruned with a factor of 10% (figure 16 was created with MountainsMap®).

6. Feature parameters in ISO 25178-2

Surfaces have a wide field of application, ranging from the geosciences to the economic, social, natural and technical sciences, with surface metrology constituting a major representative of the latter. A common feature
Figure 15. (a) Original topographic surface located in the Latschur Mountains, Austria and (b) surface after the removal of twenty peaks together with their highest adjacent passes; the contour lines are spline-interpolated.

Figure 16. The effects of Wolf pruning: (a) original surface (three-dimensional view), (b) original surface (1308 motifs), (c) pruned surface with a pruning factor of 5% (387 motifs) and (d) pruned surface with a pruning factor of 10% (202 motifs); figures courtesy of F. Blateyron.
of surfaces investigated in surface metrology is that they are not atomically flat, but, on the contrary, they have pits, passes and peaks, thus looking like natural landscapes in a miniature edition, with this property justifying the application of the same analytical techniques for surface characterisation and simplification in the geosciences as well as in surface metrology. The most apparent difference, however, between topographic surfaces, analysed by GIScientists, and scale-limited surfaces, analysed by surface metrologists, is the order of magnitude; while the measuring units related to topographic surfaces are meters or kilometers, the respective units associated with scale-limited surfaces are micrometers or nanometers.

Scale-limited surfaces, which are defined in ISO 25178-2 [1], provide a flexible way for the identification of the different scales of surface texture [113] and are derived from the real surface of a workpiece in the following way: First of all, unwanted small-scale lateral components, such as measurement noise, are removed by an S-filter, with the resulting surface being termed primary surface. In a second step, the primary surface’s nominal form is eliminated by an F-Operator in order to obtain a flat surface, the so-called S-F surface. In an optional third step, unwanted large-scale lateral components are removed from an S-F surface by an L-filter, thus obtaining a so-called S-L surface. S-F surfaces and S-L surfaces are subsumed under the term scale-limited surfaces [1, 113, 114].

For the characterisation of scale-limited surfaces ISO 25178-2 [1] specifies twenty-five field parameters as well as nine feature parameters. While for the first group the whole set of points defining a scale-limited surface is considered, for the second group only a subset of surface features is taken into account (for a detailed review confer Blatéryon [110, 115]). In order to obtain one of the nine feature parameters, a five-step process consisting of (a) selection of the type of texture feature, (b) segmentation, (c) determination of the significant features, (d) selection of feature attributes and (e) quantification of feature attribute statistics has to be performed (confer figure 17) [1, 113, 114, 116], whereby the topics discussed within this paper refer primarily to steps (b) and (c) of this process.

The starting point is a scale-limited surface, whose areal features are dales and hills, whose line features are course lines and ridge lines and whose point features are pits, passes and peaks, with these features representing Morse cells of dimension two, one and zero. The basic term motif, denoting either a dale or a hill, refers to an ascending manifold or a descending manifold, respectively. It should be noted that despite the fact that the topological concepts used in ISO 25178-2 [1] and ISO 16610-85 [2] are Morse theoretic ones, none of the terms presented in section 3.2.1 is explicitly mentioned in the two standards.

As described in ISO 25178-2 [1], first of all, a watershed-induced segmentation of the scale-limited surface is performed, whose purpose is the identification of its motifs. Typically, the first segmentation always leads to an over-segmentation of the surface, which refers to the generation of a large number of motifs that are only of minor relevance with regard to surface topography. Anyway, the topological structure of the motifs (the one-to-one correspondence between the dales and the pits and between the hills and the peaks, respectively, is nowhere mentioned in the standard) is captured by a change tree as defined in section 3.2.3.

Concerning the calculation of the nine feature parameters, ISO 25178-2 [1] requires the simplification of an over-segmented surface according to formal criteria in order to achieve a more manageable segmentation. For this purpose four different criteria are specified in the standard, comprising the local pit height (peak height) as a percentage of the height difference between the highest and lowest points of the surface, the predefined circumference of a dale (hill), the area of a dale (hill) as a percentage of the definition area of the surface and the predefined volume of a dale (hill). Evidently, these criteria define the degree of simplicity of the simplified surface (confer section 5.2) and have thus to be applied as termination rule in a

![Feature characterisation according to ISO 25178-2.](image-url)
simplification procedure (confer line 3 of algorithm 1). It should be emphasised that actually only the local pit height (peak height) as a percentage of the height difference between the highest and lowest points of the surface is implemented as summarisation rule because, according to the definition of the change tree in ISO 25178-2 [1] and ISO 16610-85 [2], no other information than height is stored in it. Carr [40, 41, 96], however, has demonstrated that local geometric measures, specifying information related to the geometry of a surface, can also be stored in the change tree. These measures comprise, for instance, contour length, surface area or volume, thus corresponding to the other criteria specified in ISO 25178-2 [1].

An important element of the segmentation process is the repeated calculation of the relevance of the topological features, i.e. of the motifs, or equivalently, of the pits and the peaks of the surface (confer line 4 of algorithm 1). As already mentioned, due to the definition of the change tree in ISO 25178-2 [1] and ISO 16610-85 [2], no other information than height is stored in it. As a consequence, only the minimum of all height differences between a pit (peak) and its adjacent passes can serve as a measure for the importance of the change tree in ISO 25178-2 [1]. In a contour tree the values correspond to the weights of the edges leading from the pits and peaks to their adjacent passes. In illustrative examples, in which change trees with only few vertices and edges are displayed, the edges are often scaled in such a way that their lengths are proportional to the respective edge weights, i.e. the shorter the edge of the contour tree, the smaller the relevance of the respective motif (pit, peak). For the sake of completeness, it should be pointed out that Carr [40, 41, 96] has also described how local geometric measures may be used for guiding a simplification process.

As already mentioned, the calculation of the nine feature parameters that are specified in ISO 25178-2 [1] requires the simplification of over-segmented surfaces according to formal criteria. Algorithm 1 describes how this can be achieved by taking only topological features into account; in addition, the algorithm is formulated in such a general way that it is independent from a chosen data structure. However, as in ISO 25178-2 [1] and ISO 16610-85 [2] change trees are employed for surface representation, an adequate pruning technique has to be chosen (confer section 5.1); Wolf pruning, which is defined in ISO 16610-85 [2] and described as a stable and robust pruning technique that works well in practice [68, 116], is applied, with the algorithm specified in ISO 16610-85 [2] representing a special case of algorithm 1.

Based on the pruned change tree, or equivalently, the simplified scale-limited surface, feature attributes, i.e. measures related to the height, length, area or volume of a feature, in combination with feature attribute statistics are defined in ISO 25178-2 [1], referring to steps (d) and (e) of the process visualised in figure 17. In a last step, finally, nine feature parameters are specified, comprising the density of peaks, the arithmetic mean peak curvature, the ten-point height of surface, the five-point peak height, the five-point pit height, the mean dale area, the mean hill area, the mean dale volume and the mean hill volume.

In the event that two or more neighbouring points $x_1, x_2, \ldots, x_n$ have the same height $h$, which may happen in real-world applications according to practitioners, the surface can not be described by a Morse function anymore. This situation, however, can be avoided by changing the altitudes of the respective points subject to the relationship $h(x_i) = h + \epsilon_i$, $h(x_n) = h + \epsilon_n$, whereby each $\epsilon_i$ represents an infinitesimal small number, i.e. a tiny number in computing, and all values $\epsilon_i$ differ from each other.

For the sake of completeness, it should be pointed out that, according to Blateyron [117], in the actual version of ISO 25178-2 [1] several errors and inconsistencies can be found, whose elimination is scheduled for the forthcoming revision of the standard. The proposed modifications will refer to (a) closed and open motifs (the presently used definitions are not stable), (b) heights and feature attributes (the terms local pit height and local peak height are confusing), (c) peak curvature (the actual definition of the arithmetic mean peak curvature is misleading) and (d) shape parameters (new morphological parameters will be included).

Obviously, fractal surfaces, which are also mentioned in ISO 25178-2 [1], cannot be handled by the approach presented in this paper, as classical Morse theory applies only to smooth functions defined on $n$-dimensional manifolds. However, the question arises if these premises constitute a serious restriction, as scale-limited surfaces are not really fractal in a strict mathematical sense, because they are deliberately smoothed at some defined scale with the result that they do not exhibit the same structure at all scales [1, 114].

### 7. Conclusions

In the present paper several topological concepts exerting a great influence on surface metrology in general and the two standards ISO 25178-2 [1] and ISO 16610-85 [2] in particular were discussed. Firstly, three data structures for the topological characterisation of surfaces were presented, namely, Morse-Smale complexes, weighted surface networks and contour trees (change trees). As these data structures rest on the critical points and critical lines of a surface, it was shown, in a second step, how the relevance of these topological features can be defined appropriately. Although, in dependence of the respective data structure, different concepts exist, all of them have in common that they are based on height differences between adjacent surface-specific points. Thirdly, the Theorem of Matsumoto was presented that says, in
clear and simple terms, that, from a topological point of view, the only valid simplification of a two-dimensional surface is the pairwise elimination of a pit together with its lowest adjacent pass or of a peak together with its highest adjacent pass. Matsumoto’s Theorem, which is independent from a chosen data structure, thus states in a formal way how a single step of a simplification process of a surface has to look like. Its combination with the two concepts of relevance of a topological feature and degree of simplicity furthermore enables the specification of an algorithm for surface simplification. Finally, it was outlined how the presented mathematical concepts are included in ISO 25178-2 [1] and ISO 16610-85 [2], although many of the topological terms are not explicitly mentioned in the two standards. Precisely speaking, it was shown that (a) scale-limited surfaces can be regarded as Morse functions, that (b) change trees are applied for the topological characterisation of the scale-limited surfaces and that (c) Wolf pruning is employed for simplifying over-segmented scale-limited surfaces to enable the computation of the feature parameters that are specified in ISO 25178-2 [1].

In conclusion, some options concerning an improvement of the characterisation and simplification of scale-limited surfaces should be sketched. However, it is important to keep in mind that surfaces do not exist per se, but they represent merely models (confer section 2), whereby the user and/or engineer decides upon the features they should exhibit:

(a) Although the definitions given in ISO 25178-2 [1] and ISO 16610-85 [2] are rigorous, the question arises whether it would not be preferable to replace all definitions related to topological concepts within the two standards by such ones that are based on Morse theory in order to take advantage of a powerful mathematical tool.

(b) In section 6 it was mentioned that actually only the local pit height (peak height) as a percentage of the height difference between the highest and lowest points of the surface is implemented as termination rule for the simplification process of a surface, because no other information than height is stored in the change tree. The question arises if this drawback should not be redressed by applying one of the following two options: (i) the storage of local geometric measures within the change tree as proposed by Carr [40, 41, 96] or (ii) the application of weighted surface networks, because this data structure does not only store topological, but also geometrical information related to the lines of steepest ascent and steepest descent.

(c) As surfaces are usually represented by discrete samples of data points, a modification of the ideas described in (a) can be achieved by assuming that the respective functions are not smooth, but piecewise linear. Due to this assumption, the gradients are no longer continuous and hence they do not generate pairwise disjoint integral lines, as required for the definition of ascending and descending manifolds. Nevertheless, also for these piecewise linear manifolds an adequate theory, which goes back to Banchoff [118], exists and has already been adopted within computational geometry and computer graphics [59, 60, 63, 119–121]. In order to extend the smooth notions to piecewise linear manifolds, differential structures are applied for guiding the computations, with this method being termed the simulation of differentiability paradigm [60].

(d) Another option for a modification of the ideas presented in (a) is the application of discrete Morse theory, as introduced by Forman [122, 123], instead of smooth Morse theory. While the latter establishes a connection between the topology of a manifold and the critical points of a smooth function defined on it, Forman derived an analogue theory for functions defined on manifolds that are discretised in form of simplicial or cell complexes. In addition, Forman demonstrated how discrete analogues of intrinsically smooth concepts, such as the gradient vector field, the gradient flow or the Morse complex can be defined. As discrete Morse theory provides a robust computational framework, it is gaining more and more interest within computational geometry and computer graphics [51, 61, 62, 65, 104, 124–131].

Acknowledgments

The invitation of Richard Leach (University of Nottingham) to hold this keynote and the assistance of François Blateyron (Digital Surf) for providing the respective graphics are gratefully acknowledged.

ORCID iDs

Gert W Wolf @ https://orcid.org/0000-0001-8842-2272

References

[1] ISO 25178-2 2012 Geometrical Product Specifications (GPS)—Surface texture: areal—Part 2: Terms, definitions and surface texture parameters (Geneva: International Organization for Standardization)
[2] ISO 16610-85 2013 Geometrical Product Specifications (GPS)—Filtration—Part 85: Morphological area filters: Segmentation (Geneva: International Organization for Standardization)
[3] Stachowiak H 1973 Allgemeine Modelltheorie (Wien, New York: Springer) (https://ia800709.us.archive.org/34/items/Stachowiak1973AllgemeineModelltheorieStachowiak%20%282%1973%29%29%2A%20Allgemeine%202%20Modelltheorie.pdf) (accessed 20.02.20)

[4] Ghorpade S R and Limaye B V 2010 A Course in Multivariable Calculus and Analysis (New York: Springer) (https://doi.org/10.1007/978-1-4419-1621-0)

[5] Moskowitz M and Paliogiannis F 2011 Functions of Several Real Variables (New Jersey: World Scientific) (https://doi.org/10.1142/7672)

[6] Matsumoto Y 2002 An Introduction to Morse Theory (Providence: American Mathematical Society)

[7] Milnor J 1963 Morse Theory (Princeton: Princeton University Press)

[8] Zomorodian A J 2005 Topology for Computing (Cambridge: Cambridge University Press) (https://doi.org/10.1017/CBO9780511546945)

[9] Fomenko A T and Kunii T L 1997 Topological Modeling for Visualization (Tokyo: Springer) (https://doi.org/10.1007/978-4-431-66956-2)

[10] Gyulassy A, Pascucci V, Petkera T and Ross B 2012 The parallel computation of Morse–Smale complexes Proc. 26th Int. Parallel and Distributed Processing Symp. (IPDPS 2012) (Shanghai) pp 484–93

[11] Maxwell J C 1870 On hills and dales Dublin Phil. Mag. J. Sci. 40 421–7

[12] Böhm W, Farin G and Kahmann J 1984 A survey of curve and surface methods Comput.-Aided Geom. Des. 1 1–60

[13] Farin G 1988 Curves and Surfaces for Computer Aided Geometric Design. A Practical Guide (Boston: Academic Press)

[14] Peucker T K, Fowler R J, Little J and Mark D M 1979 The triangulated irregular network Proc. Int. Symp. Cartography and Computing: Applications in Health and Environment, vol 2 (Auto-Carto 4) (Boston) pp 96–103

[15] Puppo E, Davis L, de Menthon D and Teng Y A 1994 Parallel computation of the topology of level sets Int. J. Geom. Inf. Syst. 7 105–28

[16] Xiao X, Weiping Q, Qings Z, Yeting Z and Zhiqian D 2013 Integration method of TINs and grids for multi-resolution surface modeling Geo-Spatial Inf. Sci. 16 61–8

[17] Yang B, Shiw L and Li Q 2005 An integrated TIN and grid method for constructing multi-resolution digital terrain models Int. J. Geogr. Inf. Sci. 19 1019–38

[18] Peucker T K 1973 Geographic data structures. Progress report after one year Technical Report No. 1. Geographic Data Structures Project, ONR Contract N00014–73–C–0109, Department of Geography, Simon Fraser University

[19] Scott P J 1998 Foundations of topological characterization of surface texture Int. J. Mach. Tools Manuf. 38 559–66

[20] Scott P J 2001 An algorithm to extract critical points from lattice height data Int. J. Mach. Tools Manuf. 41 1889–97

[21] Jeong M 2014 Qualitative characteristics of fields monitored by a resource-constrained geosensor network Ph.D Thesis Department of Infrastructure Engineering, University of Melbourne.

[22] Jeong M, Duckmam H, Kealy A, Miller H J and Peiker A 2014 Decentralized and coordinate-free computation of critical points and surface networks in a discretized scalar field Int. J. Geogr. Inf. Sci. 28 1–21

[23] Chang Y C, Song G S and Hsu S K 1998 Automatic extraction of ridge and valley axes using the profile recognition and polygon-breaking algorithm Comput. Geosci. 24 853–93

[24] Fu G, Hojvat S A and Colchester A C F 2004 Integrating watersheds and critical point analysis for object detection in discrete 2D images Med. Image Anal. 8 177–85

[25] Garbrecht J and Martz L W 1997 The assignment of drainage direction over flat surfaces in raster digital elevation models J. Hydrol. 193 204–13

[26] O’Callaghan J F and Mark D M 1984 The extraction of drainage networks from digital elevation data Comput. Vis. Graph. Image Process. 28 323–44

[27] Palacios-Vélez O L, Gandy-Bernasconi W and Cuevas-Renaud B 1998 Geometric analysis of surface runoff and the computation order of unit elements in distributed hydrological models J. Hydrol. 211 266–74

[28] Saux E, Thibaud B, Li J and Kim M H 2004 A new approach for a topographic feature-based characterization of digital elevation data Proc. 12th ACM Int. Symp. Advances in Geographic Information Systems (GIS’04) (Washington) pp 73–81

[29] Haralick R M 1983 Ridges and valleys on digital images Comput. Vis. Graph. Image Process. 22 28–38

[30] Haralick R M 1984 Digital step edges from zero crossing of second directional derivatives IEEE Trans. Pattern Anal. Mach. Intell. 6 58–68

[31] Haralick R M, Watson I. T and Laffey T J 1983 The topographic primal sketch Int. J. Robot. Res. 2 50–72

[32] Zheng X, Xiong H, Gong J and Yue L 2013 A robust channel network extraction method combining discrete curve evolution and the skeleton construction technique Adv. Water Res. 83 17–27

[33] Bleau A and Leon L J 2000 Watershed-based segmentation and region merging Comput. Vis. Image Underst. 70 317–70

[34] Stoev S L and Straifer W 2000 Extracting regions of interest applying a local watershed transformation Proc. IEEE Visualization 2000 (VIS 2000) (Salt Lake City) pp 21–9

[35] Hu Y, Miller J H and Li X 2014 Detecting and analyzing mobility hotspots using surface networks Trans. GIS 18 911–35

[36] Schneider B 2005 Extraction of hierarchical surface networks from bilinear surface patches Geom. Anal. 37 24–63

[37] Schneider B and Wood J 2004 Construction of metric surface networks from raster-based DEMs Topological Data Structures for Surfaces ed S Rana (Chichester: Wiley) pp 53–70

[38] Takahashi S 2004 Algorithms for extracting surface topology from digital elevation models Topological Data Structures for Surfaces ed S Rana (Chichester: Wiley) pp 51–51

[39] 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 Comput. Graph. Forum 14 C181–92

[40] Carr H 2004 Topological manipulation of iso-surfaces PhD Thesis Department of Computer Science, University of British Columbia (https://open.library.ubc.ca/cIRcle/collections/ubc:205831/technivaluation2004/20040401/1.0051287?accountid=13672 [accessed 20.02.20])

[41] Carr H, Snoeyink J and van de Panne M 2004 Simplifying flexible iso-surfaces using local geometric measures Proc. IEEE Visualization 2004 (VIS ’04) (Austin) pp 497–504

[42] Chiang Y J, Lenz T, Lu X and Rote G 2005 Simple and optimal output-sensitive construction of contour trees using monotone paths Comput. Geom. 30 165–95

[43] Dillard S E 2009 Contour trees and cross-sections of multiphase segmentations PhD Thesis Department of Computer Science, University of California (Irvine) (https://search.proquest.com/dissertations/doi/10.2425/733210/fulltextPDF/D1AD5530049B4D3EFP?accountid=13672 [accessed 20.02.20])

[44] Pascucci V and Cole-McLaughlin K 2002 Efficient computation of the topology of level sets Proc. IEEE Visualization 2002 (VIS ’02) (Boston) pp 187–94

[45] Pascucci V and Cole-McLaughlin K 2004 Parallel computation of the topology of level sets Algorithmica 38 249–68

[46] Pascucci V, Cole-McLaughlin K and Scorzelli G 2004 Multi-resolution computation and presentation of contour trees LNLI Technical Report Number UCRL-PROC-208680 (http://pascucci.org/pdf-papers/oreremy-2003.pdf [accessed 20.02.20])

[47] Tarasov S P and Vyalii M N 1998 Construction of contour trees in 3D in O(log n) steps Proc. 14th ACM Symp. Computational Geometry (SCG98) (New York) pp 68–75

[48] Van Krevel M, van Oostrom R, Bajaj C, Pascucci V and Schikore D 1997 Contour trees and small seed sets for iso-surfaces traversal Proc. 13th ACM Symp. Computational Geometry (SCG‘97) (Nice) pp 212–20

[49] van Oostrom R W 1999 Geometric algorithms for geographic information systems PhD Thesis Informatica-instituut, Universiteit Utrecht (https://dspace.library.uu.nl/bitstream/handle/1874/847/full.pdf?sequence=1&isAllowed=y [accessed 20.02.20])
Danovaro E, de Floriani L, Magillo P, Mesmoudi M M and Gyulassy A, Bremer P T, Hamann B and Pascucci V 2008 A multi-resolution data structure for two-dimensional Morse-Smale functions Proc. IEEE Visualization 2003 (VIS’03) (Seattle) pp 139–46

Bremer P T, Weber G H, Pascucci V, Day M and Bell B J 2010 Analyzing and tracking burning structures in lean premixed hydrogen flames IEEE Trans. Vis. Comput. Graph. 16 248–60

Čomić L and de Floriani L 2008 Multi-scale 3D Morse complexes Proc. Int. Conf. Computational Sciences and its Applications 2008 (ICCSA 2008) (Perugia) pp 441–51

Čomić L and de Floriani L 2009 Tree-based encoding for cancellations on Morse complexes Combinatorial Image Analysis ed P Wiederhold and R P Barneva (Berlin: Springer) 330–43

Čomić L, de Floriani L and Iurichich F 2011 Simplifying morphological representations of 2D and 3D scalar fields Proc. 19th ACM SIGSPATIAL Int. Conf. Advances in Geographic Information Systems (GIS’11) (Chicago) pp 437–40

Danovaro E, de Floriani L, Magillo P, Mesmoudi M M and Puppo E 2003 Morphology-driven simplification and multiresolution modeling of terrains Proc. 11th ACM Int. Symp. Advances in Geographic Information Systems (GIS’03) (New Orleans) pp 63–70

Doraiswamy H, Shivasankar N, Nataraajan V and Wang Y 2013 Topological saliency Comput. Graph. 37 787–99

Edelsbrunner H, Harer J, Nataraajan V and Pascucci V 2003 Morse-Smale complexes for piecewise linear 3-manifolds Proc. 19th ACM Symp. Computational Geometry (SCG’03) (San Diego) pp 361–70

Edelsbrunner H, Harer J and Zomorodian A 2003 Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds Discrete Comput. Geom. 30 87–107

Gyulassy A G 2008 Combinatorial construction of Morse-Smale complexes for data analysis and visualization PhD Thesis Department of Computer Science, University of California, Davis (https://search.proquest.com/dissertations/docview/304646592/fulltextPDF/DE34E16737754318PQ?accountid=13672 [accessed 20.02.20])

Gyulassy A, Bremer P T, Hamann B and Pascucci V 2008 A practical approach to Morse-Smale complex computation: scalability and generality IEEE Trans. Vis. Comput. Graph. 14 1619–26

Gyulassy A, Nataraajan V, Pascucci V, Bremer P T and Hamann B 2005 Topology-based simplification for feature extraction from 3D scalar fields Proc. IEEE Visualization 2005 (VIS’05) (Minneapolis) pp 535–42

Heine C, Leitte H, Hlawitschka M, Iurichich F, de Floriani L, Scheuermann G, Hagen H and Garth C 2016 A survey of topology-based methods in visualization Comput. Graph. Forum 35 643–67

Weinkauf T and Günther D 2009 Separatrix persistence: extraction of salient edges on surfaces using topological methods Comput. Graph. Forum 28 1519–28

Scott P J 2003 Novel areal characterisation techniques Advanced Techniques for Assessment Surface Topography: Development of a Basis for 3D Surface Texture Standards ‘Surfstand’ ed L Blunt and X Jiang (London, Sterling: Kogan Page Science) pp 43–61

Scott P J 2004 An application of surface networks in surface texture Topological Data Structures for Surfaces ed S Rana (Chichester: Wiley) pp 157–66

Scott P J 2004 Pattern analysis and metrology: the extraction of stable features from observable measurements Proc. R. Soc. A 460 2845–64

Pfaltz J L 1976 Surface networks Geogr. Anal. 8 77–93

Pfaltz J L 1979 A graph grammar that describes the set of two-dimensional surface networks Graph-Grammars and Their Application to Computer Science and Biology ed V Claus, H Ehrig and G Rozenberg (Berlin: Springer) pp 379–88

Wolf G W 1988 Generalisierung topographischer Karten mittels Oberflächengraphen Diss. Thesis Department of Geography, University of Klagenfurt

Wolf G W 1988 Weighted surface networks and their application to cartographic generalisation Visualisierungstechniken und Algorithmen ed W Barth (Berlin: Springer) pp 199–212

Wolf G W 1991 A FORTRAN subroutine for cartographic generalization Comput. Geosci. 17 1359–81

Wolf G W 2004 Topographic surfaces and surface networks Topological Data Structures for Surfaces ed S Rana (Chichester: Wiley) pp 15–29

Wolf G W 2014 Knowledge diffusion from GIScience to other fields: the example of the usage of weighted surface networks in nanotechnology Int. J. Geogr. Inf. Sci. 28 1401–24

Wolf G W 2017 Scale independent surface characterisation: Geography meets precision surface metrology Precis. Eng. 49 456–80

Burdick A and Murty U S R 1978 Graph Theory with Applications (London: The Macmillan Press)

Clarke K C and Romero B E 2017 On the topology of topography: a review Cartograph Geogr. Info. Sci. 44 271–82

Okabe A and Masuyama A 2004 A method for measuring structural similarity among activity surfaces and its application to the analysis of urban population surfaces in Japan Topological Data Structures for Surfaces ed S Rana (Chichester: Wiley) pp 143–55

Rana S 2000 Experiments on the generalisation and visualisation of surface networks Working Paper Series No 24 (London: Centre for Advanced Spatial Analysis, University College) (https://discovery.ucl.ac.uk/id/eprint/1751/1/working-paper24.pdf [accessed 20.02.20])

Rana S and Dykes J 2004 Application of surface networks for augmenting the visualisation of dynamic geographic surfaces Topological Data Structures for Surfaces ed S Rana (Chichester: Wiley) pp 143–55

Rana S and Morley J 2002 Surface networks Working Paper Series No 43 (London: Centre for Advanced Spatial Analysis, University College) (https://discovery.ucl.ac.uk/id/eprint/180/1/working-paper25.pdf [accessed 20.02.20])

Sinha G and Mark D M 2010 Cognition-based extraction and modelling of topographic eminences Cartographia: Int. J. Geogr. Inf. Geovisualization 45 103–12

Wu B, Yu B, Yao S, Wu Q, Chen Z and Wu J 2019 A surface network based method for studying urban hierarchies by night time light remote sensing data Int. J. Geogr. Inf. Sci. 33 1377–98

Boyell R L and Rushton H 1963 Hybrid techniques for real-time radar simulation Proc. 1963 Fall Joint Computer Conference (AFIPS’63) (Las Vegas) pp 455–88

Boehm W B 1967 Tabular representation of multivariate functions—with applications to topographic modeling Proc. 22nd National Conf. (ACM’67) (Washington) pp 603–15

Mark D M 1977 Topological randomness of geomorphic surfaces PhD Thesis, Technical Report No. 15, Geographic Data Structures Project, ONR Contact N00014-75-C-0886 Department of Geography, Simon Fraser University

Mark D M 1979 Topology of ridge patterns: randomness and constraints Geol. Soc. Am. Bull. 90 164–72

Kweon I S and Kanade T 1991 Extracting topographic features for outdoor mobile robots Proc. 1991 IEEE Int. Conf. Robotics and Automation Volume 3 (Sacramento) pp 1992–72
Kweon I S and Kanade T 1994 Extracting topographic terrain features from elevation maps CVGIP, Image Underst. 59 171–82

Takahashi S, Takeshima Y, Nielson G M and Fujishiro I 2004 Topological volume skeletonization using adaptive tetrahedralization Proc. IEEE Conf. Geometric Modelling and Processing 2004 (GMP’04) (Beijing) pp 227–36

Svecncak A 2005 Subdomain aware contour trees and contour evolution in time–dependent scalar fields Proc. IEEE Int. Conf. Shape Modeling and Application 2005 (SMI’05) (Cambridge) pp 136–44

Bremer P T, Weber H, Tierny J, Pascucci V, Day M S and Bell J B 2011 Interactive exploration and analysis of large-scale simulations using topology-based data segmentation IEEE Trans. Vis. Comput. Graph. 17 1307–24

Wu K and Zhang S 2013 A contour tree based visualization for exploring data with uncertainty Int. J. Uncertain. Quantification 3 203–23

Carr H, Snoeyink J and van de Panne M 2010 Flexible isosurfaces: simplifying and displaying scalar topology using the contour tree Comput. Geom. 43 42–58

Sircar J K and Cebrian J A 1991 An automated approach for labeling raster digitized contour maps Photogramm. Eng. Remote Sens. 57 965–71 (https://asprs.org/wp-content/uploads/pers/1991/jul/jul_965-971.pdf) [accessed 20.02.20]

Wu C and Mackaness W 1992 Automatic contour labeling for scanned topographic maps Proc. 17th ISPRS Congress Technical Commission IV: Cartographic and Data Base Applications of Photogrammetry and Machine Vision – ISPRS Archives—Volume XXIX Part B4 (Washington) pp 265–9 (https://isprs.org/proceedings/xxix/congress/part4/265-XXIX-part4.pdf) [accessed 20.02.20]

Guilbert E 2013 Multi-level representation of terrain features on a contour map Geoinformatica 17 301–14

Guilbert E 2016 Feature-driven generalization of isobaths on nautical charts: a multi-agent system approach Trans. GIS 20 126–43

Wu Q, Liu H, Wang S, Yu B, Beck R and Hinkel K 2015 A localized contour tree method for deriving geometric and topological properties of complex surface depressions based on high-resolution topographical data Int. J. Geogr. Inf. Sci. 29 2041–60

Takahashi S, Takeshima Y and Fujishiro I 2004 Topological volume skeletonization and its application to transfer function design Graph. Models 66 24–49

Reininghaus J, Kotava N, Günther D, Kasten I, Hagen H and Hotz I 2011 A scale space based persistence measure for critical points in 2D scalar fields IEEE Trans. Vis. Comput. Graph. 17 2045–52

Gunther D 2012 Topological analysis of discrete scalar data Doctorate in Engineering Thesis Saarland University (https://doi.org/10.22028/D291.26413) [accessed 20.02.20]

Harrison J M and Lo C P 1996 PC-based two-dimensional discrete Fourier transform programs for terrain analysis Comput. Geosci. 22 419–24

Bjerke J T and Nielsen S 2003 Wavelets applied to simplification of digital terrain models Int. J. Geogr. Inf. Sci. 17 601–21

Guilbert E and Sauz E 2008 Cartographic generalisation of lines based on a B-spline snake model Int. J. Geogr. Inf. Sci. 22 847–70

Peucker T K, Fowlr R J, Little J and Mark D M 1976 Triangulated irregular networks for representing three-dimensional surfaces Technical Report No. 10, Geographic Data Structures Project, ONR Contract N00014–75–C–0886 Department of Geography, Simon Fraser University

Peucker T K, Fowlr R J, Little J and Mark D M 1978 The triangulated irregular network Proc Digital Terrain Models (DTM) Symp. Volume 1 (St. Louis) pp 516–40