

Surface Shape Understanding based on Extended Reeb Graphs

S. Biasotti, B. Falcidieno and M. Spagnuolo

Istituto di Matematica Applicata e Tecnologie Informatiche

Consiglio Nazionale delle Ricerche

Via De Marini 6 – 16149 Genova, Italy

{silvia, falcidieno, spagnuolo}@ge.imati.cnr.it

1.1. Introduction

Knowledge about the global properties of a shape and its main features is very useful for the comprehension and intelligent analysis of large data sets: the main features and their configuration are important to devise a surface understanding mechanism that discards irrelevant details without losing the overall surface structure. As far as terrain surfaces are concerned, it is also important that a description captures important topographic elements, such as peaks, pits and passes, which have a relevant semantic content and, at the same time, are formally well-defined. Critical points and their configuration, indeed, and the related theory of differential topology give a suitable framework to formalise and solve several problems related to shape understanding. Computational topology techniques provide several tools and measures for surface analysis and coding [8]: Euler's equation, Morse theory, surface networks or Reeb graphs, for example, provide highly abstract shape descriptions, with several applications to the understanding, simplification and minimal rendering of large data sets.

Obviously the best shape descriptor does not exist, and each gives a specific view of a shape. For example, surface networks give a region-oriented description of a terrain, which can be seen as decomposed in patches having their vertices at critical points, Reeb graphs, conversely, give a volume-oriented description in which hills and dales are represented explicitly together with their elevation-based adjacency relationships.

To use topological approaches in a computational context and for discrete surfaces, it is necessary to adapt to discrete surface models concepts developed for smooth manifolds, such as piecewise linear approximations. In this chapter the notion of Extended Reeb Graph (ERG) is introduced; it is based on a characterisation strategy, which defines critical points and areas by analysing the

evolution of the contour levels on a shape, including also the so-called degenerate configurations. An algorithm for the construction of the ERG extraction is also proposed.

The remainder of this chapter is organised as follows: first, an overview of the definition of critical points and Morse complexes for smooth manifolds is given; then, topological structures used for the analysis and simplification of triangular meshes are described focusing on surface networks and Reeb graphs; the characterisation, based on a surface slicing approach, and the ERG representation are presented in section 1.3; finally, in section 1.4, an algorithm for implementing the characterisation and the ERG extraction from triangular meshes is presented together with several examples; discussions and conclusions end the chapter.

1.2. Background: differential topology for surface characterisation

Theoretical approaches based on differential topology and geometry give complete answers to the problem of understanding and coding the shape of scalar fields. In general, the configuration of the critical points gives sufficient information to fully characterise the surface shape with diverse formal codings, which highlight slightly different properties of the surface. The best example is the Morse theory, which sets the background for surface networks and Reeb graphs, both being effective tools for coding the surface shape. In this section, some topological techniques for surface shape descriptions are introduced, which propose different organization and coding of the relationships among the surface features, focusing on the Reeb graph representation [18], [19].

1.2.1. Morse theory

Morse theory is a powerful tool to capture the topological structure of a shape. In fact, Morse theory states that it is possible to construct topological spaces equivalent to a given differential manifold describing the surface as a decomposition into primitive topological cells, through a limited number of information [10], [15].

Formally, let M be a smooth manifold, that is a space for which each point has a neighbourhood locally homeomorphic to the open unit ball B^n in \mathcal{R}^n , and let $f: M \rightarrow \mathcal{R}$ be a real smooth function defined on the manifold M , whose critical points are those where the gradient is zero. Then, the following definition is given:

Definition 1 (*Morse function*):

The function f is called a *Morse function* if all of its critical points are non-degenerate, where a critical point is non-degenerate if the Hessian matrix H of the second derivatives of f is non-singular at that point.

It follows that a Morse function has to be at least C^2 . Non-degenerate critical points are isolated, and, in a neighbourhood of each critical point P , the function f can be expressed in a local coordinate system as $f = f(P) - (y_1)^2 - \dots - (y_\lambda)^2 + (y_{\lambda+1})^2 + \dots + (y_n)^2$, where λ is called the index of f in P and it represents the number of negative eigenvalues of the Hessian matrix in P . Additional details can be found in [9], [10] and [15].

In the case of terrain surfaces, which are modelled by single-valued functions, the reference manifold M is a two-manifold with boundary, where all points, except those along the boundary, have a neighbourhood homeomorphic to a sphere of dimension 2, that is to a disk. Points on the boundary have a neighbourhood homeomorphic to a half-disk.

Isolevels, i.e. subsets of M having the same value of f , can also be used to describe the surface shape. Isolevels are also called contours or level sets. The topological changes in the isolevel configuration, that is contour splitting or merging, only occur in correspondence of critical points of f . In Figure 1 examples of critical points are shown together with the projection of the surface isolevels in their neighbourhood. This property can be easily extended to degenerate critical points such as the monkey saddles and, in a broader sense, to flat regions; in particular, Figure 1(c) and Figure 1(d) highlight two degenerate situations, a monkey saddle and a volcano rim respectively. In Section 1.3, we will see how the evolution of isolevels on a manifold M is used to define the Reeb graph of the manifold.

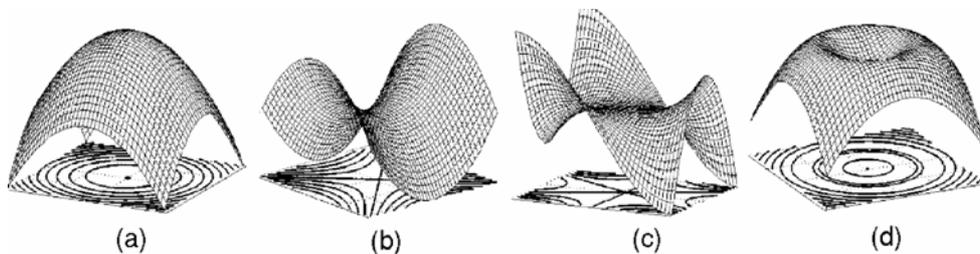


Figure 1 The behaviour of the contour levels around a maximum (a), a saddle (b), a monkey saddle (c) and a volcano rim (d).

Critical points are classified as maxima, minima and saddles, according to the behaviour of the function f around them: all the outgoing directions from a maximum (resp. minimum) point are descending (resp. ascending), while a saddle alternates at least two ascending and two descending directions.

In addition, given a Morse function f , a smooth manifold without boundary satisfies the so-called Euler formula, which states that the number of non-degenerate maximum (M), saddle (p) and minimum (m) points verifies the relation $M - p + m = 2(1-g) = \chi$, where g represents the genus of the surface and χ is called the Euler characteristics of the surface. However, considering the right

contribution of each critical point, this relation can be extended to the degenerate ones, as shown in [1] and [5].

Among all the possible Morse functions, the height function, that associates to each surface point its elevation, may be effectively used to study the surface shape in the Euclidean space. In particular, the level sets of a height function associated to a surface are the intersections of the surface with planes orthogonal to a given direction.

In [3] Banchoff presented a full framework which may be regarded as the discrete counterpart of the Morse theory, where critical points and their relationships are formally defined for triangle meshes. A basic assumption of this approach and its derived applications [2] and [11] concerns the behaviour of the scalar field at the vertices of the triangle mesh, since adjacent vertices, i.e. vertices joined by an edge, are required to have different field values. This hypothesis is needed to avoid the typical problem represented by degenerate critical points, that is non-isolated critical points such as plateaux and flat areas of the surface. Methods proposed in the literature usually do not consider the problem, delegating the solution of problematic cases to local adjustments or perturbations. This strategy, however, while solving theoretically the problem can lead to a wrong interpretation of the shape by introducing artefacts, which do not correspond to any shape feature. Also, many of the proposed computational approaches suffer from numerical instability since many degeneracies occur in real situations.

1.2.2. Surface networks

As Maxwell already guessed [14], critical points play a fundamental role for fully understanding the global topology of a shape. Topological networks, which code the relationships among the critical points, have been extensively studied; in particular surface networks have been proposed by Pfaltz [17] for the analysis of geographical surfaces. Such structures code in a graph the relation among the critical points of a surface, which are joined in the structure if there is an *integral curve* connecting them, i.e. a curve everywhere tangent to the gradient vector field. Integral curves originate from a critical point and flow to another critical point, or boundary component, and follow the maximum increasing growth of the height function, hence they cannot be closed (nor infinite) and do not intersect each other except at the critical points. In practice integral curves originate from each minimum in every directions and converge either to a saddle or a maximum, while only a finite number of integral curves can start from a saddle point.

Nackman in [16] introduced the idea of *critical point configuration graph*. Under the hypothesis the height function is Morse he demonstrated that a surface network can assume only a finite number of configurations on the surface, which induce a surface subdivision into zones of constant

first derivative behaviour, the so-called *slope-districts*. In particular, the slope districts are classified into four classes only. Then, the surface networks can be represented through a limited number of primitives, whose nodes are the critical points and whose arcs are detected through the steepest ascending directions on the surface.

For applications of the surface network framework to the GIS context see this book, part III.

1.2.3. The Reeb graph

In this chapter we are focusing on the approach proposed by Reeb to code the evolution and the arrangement of isolevel curves [18]. In the general case, the Reeb graph of a manifold M under a mapping function f is defined as follows.

Definition 2 (Reeb graph)

Let $f : M \rightarrow \mathfrak{R}$ be a real valued function on a compact manifold M . The Reeb graph of M with respect to f is the quotient space of $M \times \mathfrak{R}$ defined by the equivalence relation ‘ \sim ’ given by:

$$(X_1, f(X_1)) \sim (X_2, f(X_2)) \Leftrightarrow f(X_1) = f(X_2) \text{ and } X_1 \text{ and } X_2 \text{ are in the same connected component of } f^{-1}(f(X_1)).$$

Therefore, the Reeb graph of M collapses into one element all points having the same value under the real function f and being in the same connected component. Moreover, since the topological changes of the level sets occur only in correspondence of critical points, the Reeb quotient space can be effectively represented as a graph structure: a node is defined for each critical level of f , which corresponds to the creation, merging, split or deletion of a contour, that is, to topological changes affecting the number of connected components in the counterimage of f ; at each node, a number of arcs is defined corresponding to the number of connected components of the counterimage of f , each joining two successive critical levels in their own component. If an arc joins two nodes, n_1 and n_2 , then the topology of isolevels on M between the height levels n_1 and n_2 does not change along the connected component of M joining the corresponding critical points.

Therefore, the Reeb graph of M under the height function f can be defined as $RG_f(M) = (P_f(M), A_f(M))$, where the node set is defined by $P_f(M) = \{P_i \in M, P_i \text{ is a critical point of } f(M)\}$ and the arc set $A_f(M)$ is defined as stated before.

The arcs of $RG_f(M)$ can be oriented according to the increasing value of the height function f , that is, if $a = (n_1, n_2)$ is an arc of the graph, then $f(n_1) < f(n_2)$. Since the arcs of $RG_f(M)$ are oriented, none oriented path of $RG_f(M)$ can start and end at the same node, hence the Reeb graph is a-cyclic. Moreover, if f is Morse, the nodes have at most degree three.

With regard to terrain surfaces, these are mathematically modelled as scalar fields $h: D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ such that $h:(x,y) \rightarrow z = \text{height}(x,y)$. In this case, the manifold is defined by the points in $M = \{P \in \mathbb{R}^3 / P = (x,y,h(x,y))\}$ and the height function f is naturally defined over M as $f(P): M \rightarrow \mathbb{R}$ such that $f(P) = f((x,y,h(x,y))) = h(x,y)$. Terrain surfaces are therefore represented by scalar fields with boundary, but the Reeb graph can be always defined by adding a *minimum* to the set of critical points, which virtually closes the surface and makes it homeomorphic to a sphere, as shown in [4], [20] and [22]. Reeb graphs of terrain surfaces can be always represented as trees, where the root is given by this virtual closure of the surface.

The Reeb graph of a terrain surface M , under its natural height function, codes the shape of M in terms of the critical points of f , which are associated to meaningful topographic features, i.e. peaks, pits or passes, structured into a topologically consistent framework.

In Figure 2(a) the points drawn on the manifold represent the equivalence classes of an elementary terrain surface with respect to the height function. In Figure 2(b) the Reeb's quotient space is represented as a *traditional* graph where the equivalence classes are grouped into arcs.

Since the choice of the height function depends on the surface embedding, a manifold admits different Reeb graphs; however, this is not a problem for terrain surfaces which have a natural privileged direction.

Since the Reeb graph is not limited to scalar fields but it is really useful for analysing surfaces of arbitrary topology, it might be also extended to represent more general terrain surfaces having also vertical walls or cavities.

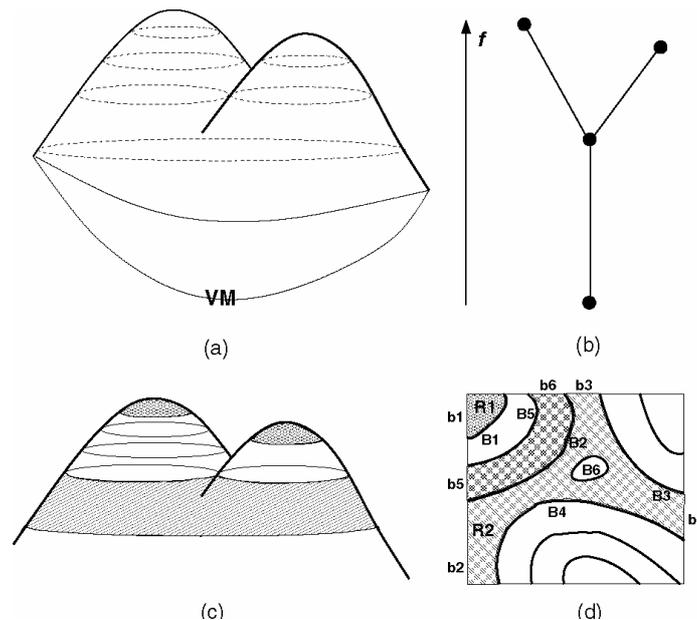


Figure 2 Reeb equivalence classes (dotted lines) (a) and Reeb graph (b) of a simple surface. The introduction of a virtual minimum makes the surface topologically equivalent to a sphere. The dark regions in (c) are critical areas, the white are the ‘regular’ ones. In (d) the regions R1, R2 and R3 and their boundary components are

highlighted; the capital labels indicate the contours of $C(M)$ and the small ones are portions of the surface boundary.

1.3. Generalised surface characterisation

As shown in section 1.2, knowledge about critical points is crucial for understanding and organizing the topological structure of a surface. Unfortunately the hypothesis that a surface is only continuous does not guarantee that the associated height function is Morse, neither derivable. Moreover, it would be desirable to distinguish among small details and relevant features of the surface, especially when dealing with rough surfaces as terrains. Many of the existing approaches to the characterisation of discrete surfaces use local point-wise criteria to detect and classify critical points: for example, triangle meshes are analysed in [3], [7], [11] and [20] by checking the height difference between a vertex and the adjacent ones in its star-neighbourhood, and by producing a topological coding, which is an adaptation of the surface network structure to piecewise-linear surfaces. Two drawbacks can be identified: first, these methods rely on the hypothesis that all edge-adjacent vertices have different height; second, the number of the detected critical points is usually very high and pruning or simplification steps are necessary to make the resulting structures understandable.

Our aim is to faithfully represent the surface topology and shape, without any height shift at surface vertices, by using an extended characterisation, which can handle degenerate as well as non-simple critical points and can be tuned to filter *small* features. Our approach is based on the use of contours for characterising the surface shape and constructing a topological structure, the Extended Reeb Graph, which represents the configuration of the critical areas of the surface. This extended characterisation is a generalisation of our previous work, see [4] and [5], in terms of both characterisation definition and algorithm for the extraction of the Reeb graph. Our approach is also similar to the method proposed in [13] for supporting the computation of intersections between parametric surfaces.

1.1.1. Definition of critical areas

A terrain surface M is characterised by sweeping slicing planes along the height direction and analysing the configuration and topological changes of the resulting isolevels, or contours. These contours decompose M into a set of regions, whose boundaries contain complete information for detecting critical areas and for classifying them as maximum, minimum and saddle areas. For example, if a contour does not contain any other contours and its elevation is higher than the successive one, then it identifies a maximum area. Our generalised characterisation corresponds to the localization of these critical areas on M , aimed at region-oriented rather than point-oriented

classification of the behaviour of M . All subsets of M defined by counterimages of critical values of f will be considered *critical areas* of M and they can be points, lines and regions.

Since terrain surfaces are surfaces with one boundary, it is also necessary to give a unique interpretation of the critical points on the boundary. This is achieved by the insertion of a global virtual minimum point, so that the outgoing directions from the surface boundary are only descending and M is virtually closed.

As shown in section 1.2 there is a tight correspondence between the existence of critical points, or areas, and the evolution of the height contours on the surface. The use of height contours has also an inherent and efficient filtering effect, which is related to the frequency or distribution of the slicing planes.

While the filtering effect will be discussed later in this section, we will assume for now that the variation interval $[f_{min}, f_{max}]$ of the height function is uniformly sliced with np planes, at a distance dp between them. The relationship between np and dp is: $np = (f_{max} - f_{min})/dp$, and the first plane is located at the height value $f_{min} + dp/2$. Moreover, we will consider that all contours are non degenerate, that is the slicing planes are never tangent to M . Details on the implementation aspects are given in section 1.4. Let $C(M)$ be the set of the resulting contour levels of the surface M , without any specific ordering. Each contour is either a simple closed line or an open line with the end points on the surface boundary B_M .

The contours in $C(M)$ fully decompose the surface M into sub-regions, which correspond either to *critical* or *regular* areas. Let $B_M(R)$ be the boundary of a region R and bb the number of its connected components; in general a connected component of $B_M(R)$ may be either a closed contour, or it may be composed by a connected and closed sequence of open contour lines and B_M parts. Note that in this latter case, if this type of component exists, then it is only one corresponding to the external boundary component of the region R . Therefore, the boundary of a region R on M is defined by $B_M(R) = B_1 \cup B_2 \cup \dots \cup B_n \cup b_1 \cup \dots \cup b_k$ where $B_i \in C(M)$ and each b_j is a portion of the surface boundary, B_M . Obviously, the boundary components $b_1 \cup b_2 \cup \dots \cup b_k$ are missing when the region does not intersect B_M , that is, the sub-region R is fully contained within the surface domain.

According to the definition of contours, if an element of $C(M)$ intersects a region R then it has to be completely part of its boundary $B_R(M)$. If the region R intersects the surface boundary B_M then the external component of $B_R(M)$ is a closed sequence of open contours connected among them through b_j components, as shown in Figure 2(d). With reference to Figure 2(d), the boundary components of \mathbf{R}_2 are made of the ordered sequence union $\mathbf{b}_2, \mathbf{B}_4, \mathbf{b}_4, \mathbf{B}_3, \mathbf{b}_3, \mathbf{B}_2$ and the boundary component \mathbf{B}_6 ; in this case bb is equal to two. In particular, with reference to the region \mathbf{R}_2 the B_i components correspond to $\mathbf{B}_2, \mathbf{B}_3, \mathbf{B}_4$ and \mathbf{B}_6 , while the b_j ones are given by $\mathbf{b}_2, \mathbf{b}_3$ and \mathbf{b}_4 .

A generic region R of M is classified according to the number and behaviour of its boundary components. Since the interior of any region R is well-defined, it is possible to associate so-called *outgoing directions* to each component of $B_R(M)$, which are needed to classify the region type. In particular, to all closed components of $B_R(M)$ only one outgoing direction is associated, while to the component intersecting B_M , if any, one outgoing direction is associated to each composing part. Each outgoing direction is classified as ascending or descending according to the behaviour of f across the corresponding boundary component. If the f value decreases (resp. increases) walking from the inside towards the outside of the region through the boundary component B_i , then the associated outgoing direction is descending (resp. ascending). The existence of the virtual minimum, indeed, does not alter the surface characterisation but implies that, during the classification process, each boundary component b_j has to be considered as a descending direction.

Given a region R and its boundary $B_R(M)$, the following classification scheme is adopted:

- R is a **maximum** area iff all the outgoing directions from $B_R(M)$ are descending, see Figure 3;
- R is a **minimum** area iff all the outgoing directions from $B_R(M)$ are ascending and $B_R(M)$ does not intersect the surface boundary, that is, $k=0$, see Figure 3(c);
- R is a **saddle** area iff either $k=0$, $bb>2$ and there are both ascending and descending outgoing directions from $B_R(M)$, or $k>0$ and $B_R(M)$ verifies at least one of the following conditions, (see Figure 3(a,b)):
 - a. $bb=1$ and there are at least two ascending outgoing directions;
 - b. $bb>1$ and at least one of the open boundary components $B_i \in B_R(M)$ has an outgoing ascending direction;
- finally, R is called **regular** iff it does not belong to the previous categories, see Figure 2(c).

With reference to Figure 2(c), the dark regions represent three critical areas, while the white ones correspond to regular areas. In addition to the previous classification scheme, a further distinction between simple and multi-connected minimum and maximum areas is done: **simple** critical areas are minima (resp. maxima) that correspond to a simply-connected region and **complex** the other ones. Moreover, due to the assumption that all the outgoing directions across the surface boundary B_M are descending, minima cannot be adjacent to B_M , and in this sense the classification of minima and maxima is not fully symmetrical. In particular, the dark regions of the image in Figure 3(a) represent critical areas. which do not belong to the boundary surface, while the regions in Figure 3(b) do.

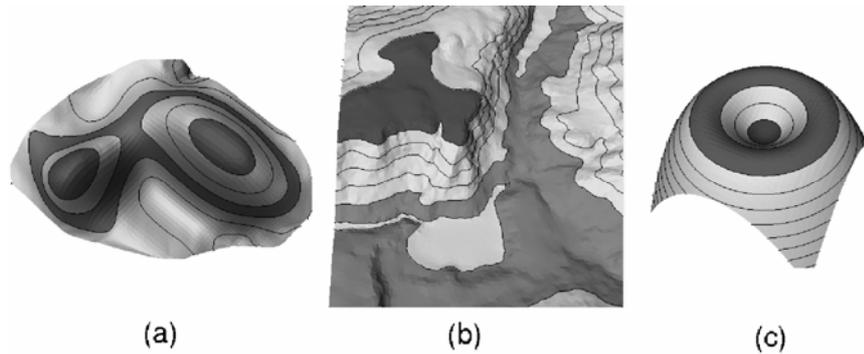


Figure 3 Maximum and saddle characterisation for regions non-intersecting (a) and intersecting the surface boundary. In (c) a minimum and a non-simply connected maximum are presented.

Let us now discuss the relation between the distribution of slicing planes and the size of the features detected. First of all, for terrain surfaces, the notion of *size* can be associated only to maximum and minimum areas, either simple or complex, and it corresponds to the height difference between the critical level and the closest adjacent saddle level. The adopted uniform slicing guarantees that all features having size greater than dp are detected. Features whose size is less than dp are discarded, except those that extend across a slicing plane. To make the filtering effect homogeneous, the contour behaviour is re-computed at a distance dp from the point q in the critical area, which has the maximum height variation within the region. In Figure 4 an example is given: the size of the feature (h) is smaller than dp and the maximum q disappears when the contour level c_1 is replaced by c_2 . In this way all the features having size greater than dp are recognised and the smaller ones are discarded.

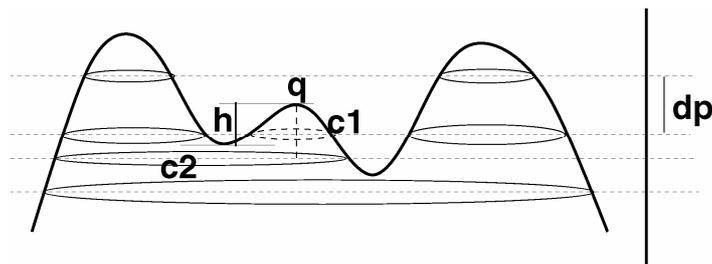


Figure 4 The feature in the middle has size h which is less than the slicing step dp , hence it is discarded during the characterisation process.

1.3.1. From critical areas to the Extended Reeb graph

The generalised characterisation just described can be coded as an Extended Reeb Graph by simply extending the equivalence relation used in the Reeb graph. Let $f: M^* \rightarrow \mathbb{R}$ be the height function defined on the virtual closing M^* of the surface M , and let $[f_{min}, f_{max}]$ be an interval containing the

variation interval of f on the surface M , and $f_{min} < f_1 < \dots < f_h < f_{max}$ the height distribution of the contour levels $C(M)$, which are supposed to be all non degenerate contours. We observe that the relations $f_{min} < f_1$ and $f_h < f_{max}$ holds, because if $f_{min}=f_1$ and $f_h=f_{max}$ the horizontal planes would be somewhere tangent to M and some contours would be degenerate. In addition, let $I=\{(f_{min}, f_1), (f_i, f_{i+1}), i=1\dots h-1, \text{ and } (f_h, f_{max})\} \cup \{f_{min}, f_1, \dots, f_h, f_{max}\}$ be the partition of the interval $[f_{min}, f_{max}]$ provided by the set of the $h+1$ interior parts of $[f_{min}, f_1, \dots, f_h, f_{max}]$ and the height values of the contour levels.

Definition 3

An *extended Reeb equivalence* between two points $P, Q \in M^*$ is given by the following conditions:

- $f(P), f(Q)$ belong to the same element of $t \in I$;
- P and Q belong to the same connected component of $f^{-1}(f(t))$, $t \in I$.

Therefore, by applying the notion of the quotient relation in definition 3, it follows that all the points belonging to a region R are Reeb-equivalent in the extended sense and they may therefore collapse into the same point of the quotient space. The quotient space obtained from such a relation is called *Extended Reeb (ER) quotient space*. Moreover the ER quotient space, which is an abstract sub-space of M^* and is independent from the geometry, may be represented as a traditional graph which is called the *Extended Reeb Graph (ERG)*.

To represent the ER quotient space as a graph, the classes which are defined by points on contours are represented by *connecting points*, while all other classes are represented by normal points, simply called points. Connecting points are representative of contours and normal points are representative of regions. A point p representing a region R is adjacent through a *connecting point* to another point q representing another region R' in the quotient space, and a normal point is adjacent to as many connecting points as the number of connected components of the boundary of the associated region. From this point of view, the image of a regular region of M^* in the ER quotient space is adjacent only to two connecting points. Therefore, the connectivity changes of the graph representation are concentrated in the image of the critical areas, and they are equivalent to the standard Reeb graph representation which can be easily derived by merging the intermediate nodes representing regular areas into a single arc. After this merging step, the ERG simply consists of nodes representing critical areas and the associated connecting arcs.

Finally, in the Reeb representation complex areas are distinguished from simple ones by labelling the graph nodes as *macro-nodes* in the former case, and *nodes* in the latter one; that is the macro-nodes are those particular leaf nodes with only ingoing or, respectively, outgoing arcs and whose degree is at least two.

Starting from the surface characterisation previously defined and considering the introduction of the global virtual minimum, V_M , the relationship among the critical points expressed in the Euler formula may be recovered also for the critical areas, as shown in [1] and [5]. The generalised Euler formula has to take into account the number of simple as well as complex critical areas. For each complex critical area, c_a , we consider the number $mc_a = i_b - 1$, where i_b represents the number of inner boundary components of c_a . Then, if P_{mc} is the sum of all the contributions of the complex areas, the Euler formula in section 1.2 becomes $M - p + m - P_{mc} + VM = \chi$. The contribution of the i -th critical area is provided by $2 - bb_i$, where bb_i is the number of its boundary components and the Euler relation: $V_M + \sum(2 - bb_i) = \chi$. Because the number of boundary components of such a critical area corresponds to the degree of the node in the Reeb graph G , the previous relation can be rewritten as $\sum(2 - \delta_i) = \chi - V_M$, where δ_i is the degree of the i -th node of G . Considering that the sum of all the node degrees is twice the number of arcs E of G (as each arc is computed in the sum for two nodes) and the contribution of V_M is one, the previous relation can be further expressed by: $2(N - E) = \chi - 1$, where N represents the number of critical areas of M .

1.4. ERG extraction

As shown in section 1.3, the quotient space defined by the extended Reeb equivalence relation can be represented in terms of a graph. Through the extended definition of critical areas proposed in section 1.3.1, the application domain can be extended to generic continuous surfaces, without any artefacts [4]. Then, the approach proposed in this chapter is actually not an extension of the Reeb graph itself, but rather a full application of its definition in the discrete domain, which does not require the height function to be Morse.

In this section, a short description of the algorithm for characterising a triangle mesh is given. The extraction and classification of critical areas is done first by computing and inserting a suitable number of contours into the triangle mesh, and second by reconstructing and classifying the boundaries of the regions delimited by the inserted contours, according to the scheme proposed in section 1.3.1.

The computation and the insertion of the contours into the mesh is done in a single step. The contour levels $C(M)$ inserted into the mesh model are used as constraints for the region detection process, which uses a region-growing strategy. The insertion of a contour C into M is computed as follows: given a slicing plane π , a seed point $p \in C$ is computed by selecting an edge e , which properly intersects π , that is e does not belong to π nor intersects it in a vertex. C is extracted by starting from p and moving horizontally by adjacency on the mesh until either p or the surface boundary is reached. If the surface boundary has been reached, C is an open contour and the

algorithm restarts from p in the opposite direction until the surface boundary is reached again. If the points of C are not vertices of the mesh, they are inserted into the mesh. The mesh is locally re-triangulated in order to obtain a valid mesh, and the contours are inserted into the mesh as constrained edges. This process stops when all the planes have been considered. This procedure guarantees that degenerated contours as points, lines, etc. are not taken into account. Then, the intersections of the model with the slicing planes are computed and stored as a set of connected components, which can be also open, in correspondence of the surface boundary intersection.

The insertion of $C(M)$ decomposes the triangle mesh into a set of regions, each bounded by $C(M)$ elements and mesh boundary edges. These regions are detected by labelling all triangles in the mesh, with a region-growing process which propagates the label from a triangle to its adjacent ones without crossing any constraint. At the end of this labelling phase, all triangles having the same label identify a region. Then, the boundary of each region is detected and the associated outgoing directions are classified. Starting from any edge of the region boundary, the associated connected component is fully traced using edge-vertex adjacency. If the component is closed, then there is only one outgoing direction, which can be easily classified by checking the elevation of any vertex inside the adjacent region. If the traced component is open, then the tracing has to continue also along the mesh boundary, and the whole component will consist of a sequence of open contours and boundary parts. The tracing can be done since all triangles are labelled with the region label. In this case, each part of the boundary component defines an outgoing direction which has to be classified. Finally, the number of boundary components bb and their classification allow distinguishing between simple and complex critical areas.

According to the graph representation of the extended Reeb's quotient space, each node of the graph corresponds to a critical area; in particular, when the critical region recognised as a maximum/minimum area is complex, a macro-node is defined, with as many arcs as the inner components of the critical region. Since each arc corresponds to a connected component of the manifold between two critical areas, the Reeb graph extraction is based on tracking the evolution of contour lines.

When the critical areas have been recognised, the *ERG* is initialised by creating the node corresponding to the virtual minimum, V_M . The V_M is connected to the saddle having the minimum elevation and external to each macro-node. If such a saddle does not exist, the V_M is connected to the nearest (in terms of geodesic distance) complex maximum area, otherwise, if there are not complex maxima, the *ERG* is a trivial graph connecting the V_M to the only simple maximum existing and the surface is topologically equivalent to a sphere [15].

Our algorithm for the extraction of the ERG runs in two steps: first, the arcs between minima (resp. maxima) and saddles are inserted, then the other ones are detected. In the following, a construction algorithm is described using a C pseudo-code:

```

ERG_Construction(N,A)
    /*The ERG is defined by the set of nodes, N, and of arcs, A*/
    { N=CriticalAreasRecognition(tin, contours);
        /* Identify critical areas and initialise the virtual minimum */
    OrderAreas(N);          /* Order the Critical Areas by elevation */
    ConnectVirtualMinimum(N); /*Create a virtual minimum and connect it
                                to the node the most appropriate */
    ExpandMaxima&Minima(N,A); /*Leaf arc extraction */
    for (each node in N)
        {if (IsGrowingArea(node))
            {for ( each non visited growing direction node)
                {while ((not(findBoundarySurface)) or (not(findOtherCriticalArea)))
                    ExpandToUpperLevel(node);
                    if (R=OtherAreaReached)
                        ConnectWithArc(node, R);
                }
            }
        }
    }

```

The function ‘*ExpandMaxima&Minima(N)*’ connects all the maxima and minima to their nearest (in terms of region expansion) critical area and extracts a subset of Reeb arcs, while the function ‘*IsGrowingArea(node)*’ returns a Boolean value, which is *TRUE* if the critical area has at least one growing direction that has not been visited yet. In Figure 5 the main steps of the *ERG* extraction process are depicted; Figure 5(a) represents how the maxima (resp. minima) are expanded until other critical areas are reached and the corresponding graph representation, while Figure 5(b) shows how the algorithm works for completing the area expansion process.

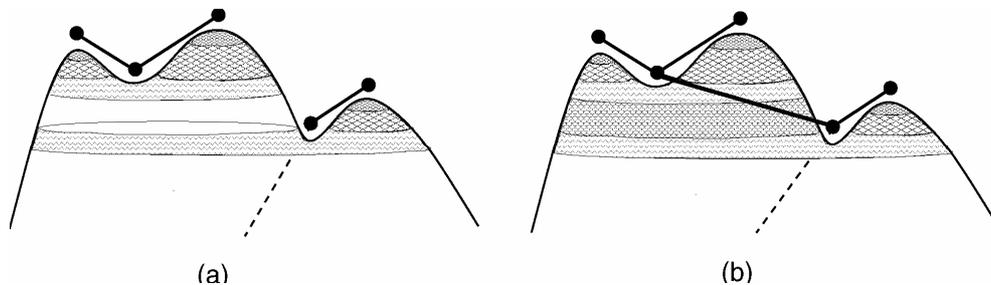


Figure 5 Two steps in the pipeline of the ERG extraction.

Some results of our *ERG* extraction for real terrains are provided in Figure 6. The nodes of the *ERG* representation are coloured according to the meaning of the corresponding critical areas into the models. In particular, the maxima are depicted in red, the minima in blue and the saddles in green, while the virtual minimum is represented in yellow. Moreover, in Figure 6 we show the simplified models obtained considering only the mesh vertices, which form the boundary of all the critical areas of the models. The original models of Figure 6(a,c) have 160000 and 129600 vertices, respectively, while the simplified ones in Figure 7 have resp. 19200 and 26200 vertices; it is important to point out that the simplification provided by the *ERG* mainly depends on the *topological* complexity of the models rather than on the number of the original vertices.

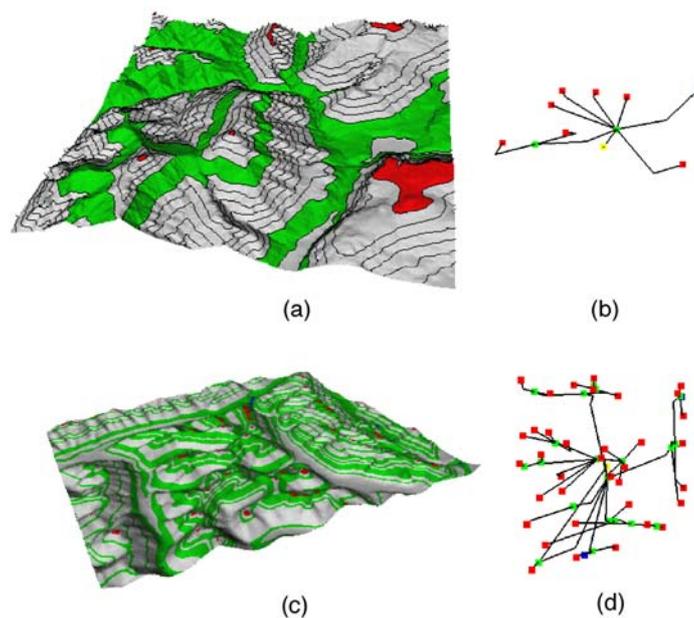


Figure 6 Two terrain models (a) and (c) and their Reeb graph representations (b) and (d). The models in (a) and (c) are freely available at <http://www.geographx.co.nz/>.

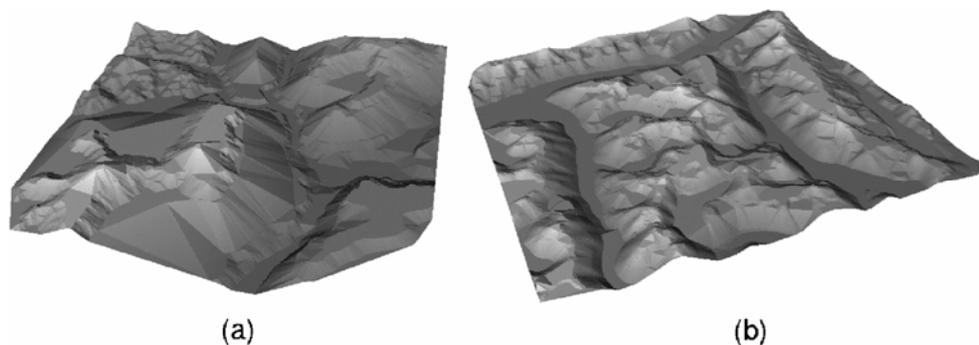


Figure 7 Examples of simplification obtained by considering only the boundaries of critical areas. In (a) the simplified model of the terrain given in Figure 6(a), and in (b) that of Figure 6(c).

1.4.1. Computational complexity

The computational cost of the whole algorithm for the ERG extraction is given by the sum of the cost of its single subparts, that is the insertion of contour levels into the mesh, the extraction of the critical areas and the final expansion process.

Given the surface mesh, the insertion of the contour levels $C(M)$ depends on both the number of vertices of the original triangulation, n , and the number m of the vertices of $C(M)$. Because the number of edges and triangles has the same order as the number of vertices, checking the edge-to-plane intersection requires $O(\max(m, n \log(n)))$ operations. In fact, the edges of the mesh are sorted in $O(n \log(n))$ operations, while $O(\max(m, n))$ is the number of intersection tests. Finally the insertion of the whole set of constraints requires $O(m)$ edge splits.

With regard to the computational complexity of the characterisation process, the recognition of critical areas is linear in the number of mesh triangles, then it requires $O(n+m)$ operations, because the number of triangles in the constrained mesh has the same order of the sum of original vertices and the constrained ones. Also during the arc completion step, the triangles are processed once and the complexity still is $O(n+m)$, so that the total computational cost of the ERG extraction mainly depends on the insertion of contours into the mesh. Therefore, the whole process, starting from a generic triangulation, requires $O(\max(m+n, n \log(n)))$ operations. Finally, we observe that, if we consider a generic triangle mesh, the average size of m is $O(n \log(n))$ even if in the worst case, m could be $O(n^2)$.

1.5. Discussion and final remarks

The generalised characterisation and the ERG coding provide a compact representation of the main features of a terrain surface, which is effectively represented as a configuration of hills and dales.

With regard to the feature extraction step, the mesh characterisation based on the classical height comparison at mesh vertices, as classically proposed in [3], [7], [11] and [20], can be recovered also through our method. It is sufficient, indeed, to slice the mesh in correspondence of the midpoint of each edge; in this way all the original mesh vertices would lie in a separate region and the characterisation obtained through the mesh contouring would be equivalent to consider the star region of each vertex.

Finding the best compromise between the effectiveness of feature extraction and the number of slicing planes is the most critical point of the method. A first solution is to characterise the mesh as proposed in [7] and [11], by slicing the mesh with planes placed at optimal positions: one plane directly below (resp. above) maxima (resp. minima), and two planes for saddles, one above and one

below. In this case the number of slicing planes considerably decreases but the number of features does not, and the results would still be sensitive to small variations of the vertex elevation.

Using the uniform slicing, the surface shape is described by the topological coding of its features at a fixed resolution dp . In many cases, however, a description at different scales could be more effective. This could be achieved by adopting a multi-resolution slicing process of the mesh as proposed in [12]: a sequence of Reeb graphs can be extracted by halving the distance interval between the slicing planes until a threshold defined by the user is reached. At each step, new nodes and arcs might be inserted into the graph as shown in Figure 8, but there is a hierarchical relation between the nodes of the current graph and the previous one [1].

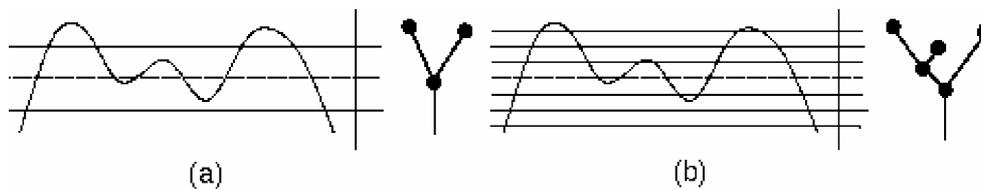


Figure 8 Reeb graph variation when halving the distance among the sections.

In our setting, a multi-resolution ERG extraction can be implemented by iteratively halving the height interval $[f_{min}, f_{max}]$; for example the graphs proposed in Figure 6 have been obtained with 32 subdivisions of the interval $[f_{min}, f_{max}]$. The power of this approach is clear: the surface shape can be processed at different levels of detail and the estimation of its features is automatically provided.

In addition, we notice that adopting the mesh characterisation approach based on the neighbours of each vertex, the Reeb graph is equivalent to that provided by the contour tree, as proposed in [6] and [21]. In fact, both structures have a common root in Maxwell's paper [14] and pursue the aim of organizing the contour levels of a two-dimensional surface in a systematic and topologically correct way. However, the contour trees have been proposed only for scalar fields, while the Reeb graphs have been studied for generic two-manifold and successfully applied to arbitrary complex surfaces; as an example our approach works also on terrain surfaces with vertical walls and cavities.

Considering simple Morse functions, i.e. functions whose critical points are non-degenerate and not at the same level, , Reeb graphs and surface networks may be easily compared: the Reeb graph is a subgraph of the surface network, at least for the arcs not involving the boundary. An algorithm for the extraction of Reeb graphs from surface networks has been, for example, proposed in [20]. Both graphs code the topological structure of a surface, with surface networks giving a surface-oriented view, while Reeb graphs giving a skeleton-like and volume-oriented description. In Figure 9 the surface network of a terrain represented by contours is compared with the corresponding Reeb

graph; all the arcs of the surface network coming from the outside of the surface boundary originate from a virtual minimum, which is depicted for the Reeb graph structure.

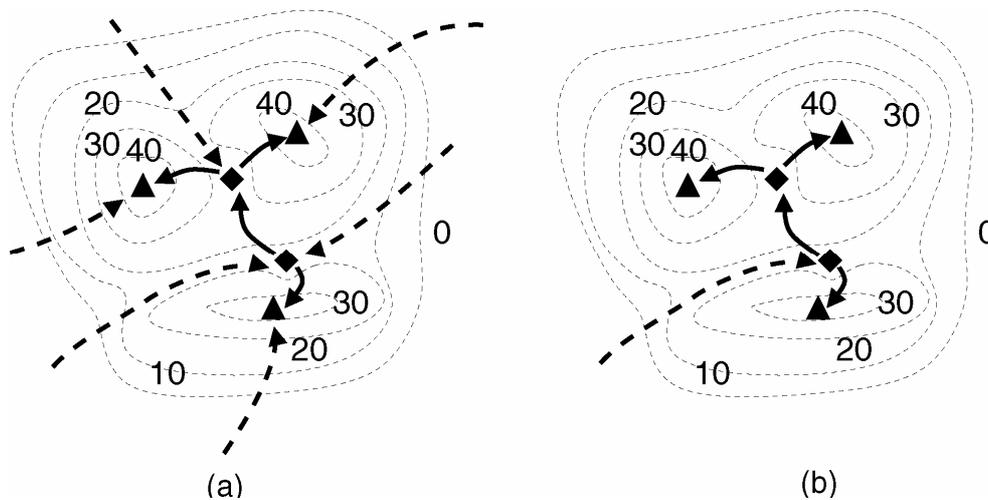


Figure 9 The surface network structure (a) and the Reeb graph (b) of the same terrain model.

In the generalised version presented in this chapter, surface networks and ERG cannot be directly compared. Surface networks obviously fail if degenerate critical points exist, and, to our knowledge, there is no way to automatically filter the resulting features during the network delineation process. Conversely, the ERG construction process is stable and it provides a simplified configuration of the terrain features, which easily and efficiently supports the minimal rendering of large terrain data.

Acknowledgements

The authors would like to thank all the people of the Shape Modelling Group at IMATI-CNR.

This work has been partially supported by the National MIUR Project ‘MACROGeo: Metodi Algoritmici e Computazionali per la Rappresentazione di Oggetti Geometrici’, FIRB grant.

References

- [1] M. Attene, S. Biasotti, M. Spagnuolo, *Shape understanding by contour driven retiling*, The Visual Computer, 2003 to appear, on-line DOI 10.1007/s00371-002-0182-y.
- [2] C. L. Bajaj, D. R. Schikore, *Topology preserving data Simplification with Error Bounds*, Computers & Graphics, Vol. 22, No. 1, pp. 3-12, 1998.
- [3] T. F. Banchoff, *Critical points and curvature for embedded polyhedral surfaces*, Am. math. Monthly, Vol. 77, pp. 475-485, 1970.

- [4] S. Biasotti, B. Falcidieno, M Spagnuolo, *Extended Reeb Graphs for Surface Understanding and Description*, in Proc. of 9th Discrete Geometry for Computer Imagery Conference, LNCS, Springer Verlag, pp. 185-197, Uppsala, 2000.
- [5] S. Biasotti, B. Falcidieno, M. Spagnuolo, *Shape Abstraction Using Computational Topology Techniques*, in 'From Geometric Modeling to Shape Modeling', (U. Cugini and M. Wozny, eds), Kluwer Academic Publishers, pp. 209-222, 2002.
- [6] H. Carr, J. Snoeyink, U. Axen, *Computing contour trees in all dimensions*, Computational Geometry, Vol. 24, pp. 75-94, 2003.
- [7] L. De Floriani, M. M. Mesmoudi, E. Danovaro, *Extraction of Critical Nets Based on a Discrete Gradient Vector Field*, In Proc. Eurographics 2002, (I.Navazo, P.Slusallek eds), pp. 373-382, Saarbrucken (Germany), 2002.
- [8] T. K. Dey, H. Edelsbrunner, S. Guha, *Computational Topology. Advances in Discrete and Computational Geometry*, eds: Chazelle B, Goodman J E, Pollack R Contemporary Mathematics 223, AMS, Providence, pp. 109-143, 1999.
- [9] H. B. Griffiths, *Surfaces*, Cambridge University Press, 1976.
- [10] V. Guillemin, A. Pollack, *Differential Topology*, Englewood Cliffs, NJ: Prentice-Hall, 1974.
- [11] H. Edelsbrunner, J. Harer, A. Zomorodian, *Hierarchical Morse complexes for piecewise linear 2-manifolds*, In Proc. 17th Sympos. Comput. Geom., pp. 70-79, 2001.
- [12] M. Hilaga, Y. Shinagawa, T. Komura and T. L. Kunii. *Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes*, ACM Computer Graphics, (Proc. of SIGGRAPH 2001), Los Angeles, 2001.
- [13] C. Jun, D. Kim, D. Kim, H. Lee, J. Hwang, T. Chang, *Surface Slicing Algorithm based on Topology Transition*, Computer Aided Design, Vol. 33, pp. 825-838, 2001.
- [14] J. C. Maxwell, *On Hills and Dales*, The London, Edinburgh and Dublin Philosophical Mag. and J. Sci. 4th Series, Vol. 40, 1870 reprinted in The Scientific Papers of James Clerk Maxwell, W. D. Niven, Ed. Vol.2, New York: Dover, 1965.
- [15] J. Milnor, *Morse Theory*, Princeton University Press, New Jersey, 1963.
- [16] L. R. Nackman, *Two-dimensional critical point configuration graphs*, IEEE Trans. Pattern Anal. Machine Intell., Vol PAMI-6 pp 442-450, July 1984.
- [17] J. L. Pfaltz, *Surface Networks*, In Geographical analysis, Vol. 8, pp. 77-93, 1990.
- [18] G. Reeb, *Sur les points singuliers d'une forme de Pfaff complètement intégrable ou d'une fonction numérique*, Comptes Rendus Acad. Sciences Paris, 222:847-849, 1946.

- [19] Y. Shinagawa, T. L. Kunii, Y. L. Kergosien, *Surface Coding Based on Morse Theory*, IEEE Computer Graphics & Applications, pp 66-78, September 1991.
- [20] S. Takahashi, T. Ikeda, Y. Shinagawa, T. L. Kunii, M. Ueda, Algorithms for Extracting Correct Critical Points and Construction Topological Graphs from Discrete geographical Elevation Data, Eurographics '95, Vol. 14, Number 3, 1995.
- [21] M. van Kreveld, R. van Oostrum, C. Bajaj, V. Pascucci, and D. Schikore, *Contour trees and small seed sets for isosurface traversal*, In Proc. 13th Annu. ACM Sympos. Comput. Geom., pp. 212--220, 1997.
- [22] J. Wood and S. Rana, *Construction of Weighted Surface Networks for the Representation and Analysis of Surface Topology*, Proceedings of the 5th International GeoComputation 2000, UK, 2000.