Reeb graphs for shape analysis and applications

S. Biasotti*, D. Giorgi, M. Spagnuolo, B. Falcidieno

IMATI - CNR, Via De Marini 6, Genova, Italy

Abstract

Reeb graphs are compact shape descriptors which convey topological information related to the level sets of a function defined on the shape. Their definition dates back to 1946, and finds its root in Morse theory. Reeb graphs as shape descriptors have been proposed to solve different problems arising in Computer Graphics, and nowadays they play a fundamental role in the field of computational topology for shape analysis. This paper provides an overview of the mathematical properties of Reeb graphs and reconstructs its history in the Computer Graphics context, with an eye towards directions of future research.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Reeb graph; Shape description; Computational topology

1. Introduction

Shape analysis and understanding are very likely research topics in Computer Graphics, with applications to many different fields, from the engineering context to medicine, entertainment, games and simulation. In the last decade, the advances in modeling, digitizing and visualizing 3D shapes have led to an explosion in the number of available 3D models, both on the Internet and in domain-specific databases. Examples are archives recording cultural heritage [1] or archives of structural data of biological macromolecules [2].

At the same time, the research interest in Computer Graphics has gradually moved from methods to represent shapes towards methods to describe shapes. As observed by Nackman, a digital model (e.g. a picture, a triangle mesh or a voxelized 3D model) is a representation which is quantitatively similar to an object, while its description is only qualitatively similar [3]. The representation of an object is detailed and accurate, but it does not explicitly contain any high-level information on the shape of the object. Conversely, the description is concise but conveys an elaborate and composite view of the object identity.

A variety of methods have been proposed in the literature to tackle the problem of shape description and reasoning. We believe that mathematics, and specifically differential topology, provides a suitable setting for formalizing and solving several problems related to shape description, analysis and understanding. The potential of this approach has been recently recognized by researchers in Computer Graphics who started a new research area called computational...
topology. This term has been introduced in [4,5] and denotes research activities involving both mathematics and computer science, in order to formalize and solve topological problems in computer applications, and to study the computational aspects of problems with a topological flavour. The idea is that many classical concepts in mathematics can be re-interpreted in a computational context, thus furnishing powerful key tools also in a discrete setting.

In particular, the classical Morse theory [6] offers a series of instruments for analyzing and describing shapes. The intuition behind Morse theory is to combine the topological exploration of a shape with quantitative measurements of its geometrical properties, provided by a real function defined on the shape. In other words, the idea is to view a shape by a pair \((S, f)\), where \(S\) is a space representing the shape and \(f\) is a real-valued mapping function defined on \(S\), and devised to identify points of interest over the shape.

The added value of approaches based on Morse theory relies on the possibility of adopting different functions as shape descriptors according to the properties and invariants that one wishes to analyze. Indeed, different choices of the function yield insights on the shape from different perspectives. For example, a terrain can be studied by analyzing the behaviour of the real function which associates with each point on the surface its height. In this case, the critical points – that is to say the points on the terrain where the gradient of the real function vanishes – are those corresponding to the intuitive notion of peaks, pits and passes, and their configuration provides an effective summary of the main shape features and structure. Changing the function, by considering, for example, the distance from a fixed point, the critical points will change accordingly and their new configuration may give insights on the shape from a different perspective.

A large number of methods rooted in this theory have been proposed in the literature [7–9], with emphasis on both mathematical and computational aspects, but the most popular descriptor is probably the Reeb graph. Originally defined in 1946 by the mathematician George Reeb [10] for smooth manifolds, the Reeb graph has been introduced in Computer Graphics by Shinagawa et al. [11] in 1991, and since then it has been proposed to solve different problems related to shape matching, morphing and coding.

The aim of this paper is to provide an overview of the mathematical properties of the Reeb graph and to reconstruct its history in the field of Computer Graphics, investigating its potential for analyzing and describing shapes.

The remainder of this paper is organized as follows. First of all, we briefly recall the essential mathematical background pertaining to Reeb graphs in Section 2. Section 3 presents the original definition of the Reeb graph, as formulated by Reeb, together with its main mathematical properties. In Section 4 we begin to detail the role of Reeb graphs in Computer Graphics; we introduce different approaches proposed in the literature to define and compute Reeb graphs and related structures in the discrete setting. Applications of Reeb graphs are described in Section 5, followed by Section 6 devoted to analyze open problems and future research trends.

2. Background notions

In this Section we report some classical results in Morse theory [12], which constitutes the essential mathematical root for Reeb graphs. For a detailed overview of Morse theory see [6].

Morse theory investigates the relation between the properties of a function defined on a manifold and the properties of the manifold itself: the aim is to derive information on the topology of the manifold from the information about the critical points of the function [6].

Intuitively, an \(n\text{-manifold} M\) is a topological Hausdorff space such that around every point there is a neighbourhood that is topologically equivalent to the open unit ball \(B^n\) in \(\mathbb{R}^n\); the number \(n\) is the dimension of the manifold [13]. This means that there exists an open cover \(\{U_i\}_{i \in A}\) of \(M\) such that for each open subset \(U_i, i \in A\), there exists a homeomorphism \(\varphi_i : U_i \rightarrow B^n\). An \(n\)-manifold with boundary is a Hausdorff topological space such that each point has an open neighbourhood homeomorphic either to the open unit ball \(B^n \subset \mathbb{R}^n\) or the open half-space \(\mathbb{R}^{n-1} \times \{x_n \in \mathbb{R} \mid x_n \geq 0\}\). Fig. 1 shows, from left to right, a 2D manifold (surface) without boundary, a surface with boundary, and a space that is not a manifold.

An important property of manifold concerns orientability. The definition of orientability relies on the properties of \(\text{atlases}\) defined on the manifold, where an \(\text{atlas} \) of \(M\) is the union of pairs \(\{(U_i, \varphi_i)\}_{i \in A}\). Given \(U_j, U_j\) two arbitrary open sets of the cover of \(M\), the function \(\varphi_{i,j} = \varphi_j \circ \varphi_i^{-1}\) restricted to the open subset \(\varphi_i(U_i \cap U_j) \subset B^n\) is a homeomorphism called a \(\text{transition function}\) or \(\text{gluing function}\) of the given atlas. Then, a manifold \(M\) is called orientable if there exists an atlas \(\{(U_i, \varphi_i)\}_{i \in A}\) on it such that the Jacobian of all transition functions \(\varphi_{i,j}\) from a chart to another is positive for all intersecting pairs of regions. Manifolds that do not satisfy this property are called
non-orientable. Depending on the type of the gluing functions (e.g., smooth, analytic, piecewise smooth, Lipschitz) the manifold is consequently named (e.g. smooth manifold, analytic manifold etc.). A compact manifold is a manifold that is compact as a topological space. A closed manifold is a compact manifold without boundary. Fig. 2 shows two classical examples of closed 2-manifolds (surfaces): the torus in Fig. 2(a) is an orientable surface, while the Klein bottle in Fig. 2(b) is a classical example of non-orientable surface.

The key to Morse theory is critical points of real functions defined on the manifold. Let $M$ be a smooth compact $n$-dimensional manifold without boundary, and $f : M \to \mathbb{R}$ a smooth function defined on it. We recall that a point $p$ of $M$ is a critical point of $f$ if all the partial derivatives of $f$ with respect to a local coordinate system about $p$ are zero in $p$. In particular, a critical point is non-degenerate if the Hessian matrix of the second partial derivatives of $f$ is non-singular at that point. Fig. 3 shows some examples of non-degenerate and degenerate critical points. For a non-degenerate critical point, the number of negative eigenvalues of the Hessian matrix of $f$ at the point determines the index of the point. Adopting the classical mathematical definition [6], we say that $f$ is a Morse function if all its critical points are non-degenerate. Notice that for applications the function $f$ is frequently required to be simple, that means all its critical points have different values (i.e. any pair $x, y$ of distinct critical points verifies $f(x) \neq f(y)$). In the recent literature, a Morse and a simple function are sometimes referred to as Morse [14].

Morse functions are everywhere dense in the space of all smooth functions on the manifold $M$. This implies that any smooth function may be transformed in a Morse function by a slight perturbation. Moreover, the set $S$ of all simple Morse functions is everywhere dense in the set of all Morse functions. As it will be discussed later on, this fact has quite an influence on the different approaches used in computational scenarios for analyzing shapes.

The properties and structure of critical points are captured by changes in the level sets and lower level sets of $M$. We begin by studying lower level sets. Setting $M_t = \{ p \in M | f(p) \leq t \} = f^{-1}((-\infty, t])$, we may wish to investigate how the lower level set $M_t$ (also known as sub-level set [15]) changes as the parameter $t$ changes (see Fig. 4(a)). Morse theory states that the topology of $M_t$ stays unchanged (formally, the homotopy type is preserved) as the parameter $t$ goes through regular values of $f$, while changes occur when $t$ passes through a critical value (see Fig. 4(b)).

In this context, an important property is that a Morse function defined on a compact manifold admits only finitely many critical points. This fact means that the homotopy of $M_t$ changes only in a finite number of times and it is
possible to define a manifold decomposition whose elements (technically called cells) correspond to the critical points of the manifold. This property is crucial, because it guarantees a finite decomposition of the manifold into simpler elements. More precisely, let \( p \) be a critical point of \( f \) with index \( \lambda \) and let \( f(p) = c \). Then, for each \( \varepsilon \) such that \( f^{-1}([c - \varepsilon, c + \varepsilon]) \) contains no critical points other than \( p \), the set \( M_{c+\varepsilon} \) has the same homotopy type of the set \( M_{c-\varepsilon} \) with a \( \lambda \)-cell attached:

\[
M_{c+\varepsilon} \cong M_{c-\varepsilon} \cup_{\varphi_p} e^\lambda.
\]

A \( \lambda \)-cell \( e^\lambda \) corresponds to the closed unit ball of dimension \( \lambda \). The attaching map \( \varphi_p \) is a continuous map from the boundary \( \partial e^\lambda \) of the \( \lambda \)-cell to the set \( M_{c-\varepsilon} \): attaching \( e^\lambda \) to \( M_{c-\varepsilon} \) by the map \( \varphi_p \) means to identify each point \( x \in \partial e^\lambda \) with the point \( \varphi_p(x) \in M_{c-\varepsilon} \) (see Fig. 4 (c)). Informally speaking, \( M \) obtained by subsequently attaching finitely many cells is a finite cell complex.

Given these results, it is now possible to state a link between the number of critical points of the function \( f \) and the global topology of the manifold \( M \), in terms of its Euler characteristic. We recall that the Euler characteristic (also known as the Euler–Poincaré characteristic), is a topological invariant defined as the alternating sum of the Betti numbers \( \beta_i \) of \( M \) [16], where \( \beta_i \) is the rank of the \( i \)th homology group of \( M \):

\[
\chi(M) = \sum_{i=0}^n (-1)^i \beta_i(M).
\]

Indeed, denoting by \( \mu_\lambda(M) \) the number of cells of dimension \( \lambda \) for each \( \lambda \), it holds

\[
\chi(M) = \sum_{\lambda=0}^n (-1)^{\lambda} \mu_\lambda.
\]

Eq. (1) is also known as the Morse–Euler formula.
Fig. 4. (a) A manifold $M$ and three lower level sets $M_a, M_b, M_c$, with respect to the height function. (b) There are no critical points in $f^{-1}([a, b])$: $M_a$ (left) and $M_b$ (right) are topologically the same. (c) The passage through the critical point $p$ of index 1 causes a topological change: $M_c$ (left) has the same homotopy type of $M_b$ with a 1-cell attached (right).

The decomposition of $M$ into cells previously described characterizes $M$ by means of the structure of a cell complex. To build the cell complex corresponding to $M$, we have to start with a single point (i.e. a 0-cell), then order and glue the remaining cells, for example, with respect to the value of $f$. Therefore there is a kind of arbitrariness in the construction of the cell complex associated with $M$. However, a more general statement claims that the cells may be attached so that their dimensions form a non-decreasing sequence [17]. In any case, as recalled by Shinagawa et al. [11] for 2-manifolds embedded in $\mathbb{R}^3$, the sequence of indices of the critical points of $f$ does not code the way the manifold is embedded into the space. This consideration has led us to consider other topological structures, such as Morse and Morse–Smale complexes or Reeb graphs. In this paper, we focus our attention on Reeb graphs, highlighting that a comparative analysis of Morse and Morse–Smale complexes and contour tree structures is addressed in [18] and a more general overview of descriptors rooted in Morse theory is proposed in [19].

The critical points of a Morse function defined on an $n$-manifold influence the behaviour not only of lower level sets, but also of level sets $V_t = f^{-1}(t)$ (also called isocontours in the case of 2-manifolds). Indeed, it happens that, when $t$ is a regular value for $f$ then $V_t$ is either empty or the union of finitely many smooth $(n-1)$-manifolds. In the particular case of surfaces (2-manifolds), it has been also demonstrated that, when the isovalue $t$ spans a range of values containing a critical value of $f$, then the level sets change; vice versa, a change in the topology of the level sets locates a critical value of $f$ [20].

The validity of Morse theory in the discrete context has received great attention in the literature, and many results have been proven [21–23]. The notion of critical points was firstly studied by Banchoff for height functions defined over polyhedral models defined over a cell decomposition of their domain. A geometric characterization of critical points was introduced taking into account the position of the tangent plane with respect to the model and the corresponding number of intersections, which define the discrete index. Within this setting Banchoff proved
that an analogous of the Morse–Euler formula (Eq. (1)) holds for general\(^1\) height functions defined on polyhedral surfaces [24].

For higher-dimensional spaces, the Betti numbers of the lower link provide a more complete characterization of discrete critical points, as suggested in [25].

The theory of Banchoff has been used by most of the authors dealing with computational topology. In many applications, however, the shapes to be analyzed are likely to have degenerate critical points. Degenerate critical points can be handled either by replacing the notion of critical point with that of critical area [26,27,23], or by perturbing and unfolding the simplicial complex [28,14].

We also mention here the existence of a different discretization approach related to Morse theory: the so-called discrete Morse theory introduced by Forman in [17,29]. The basic idea consists of analyzing a cell complex by assigning a single number to each of its cells.

Recently, the problem of constructing functions with a fixed number of critical points has been investigated [9]; indeed, Morse Theory does not guarantee that Morse functions on \( M \) have the same number of critical points or that the cell complex obtained using a given \( f \) is the “best possible”, in the sense of having the fewest number of cells. In this context, it holds that harmonic functions, which are solutions of the Laplace equation, guarantee the minimum number of local extrema, except at boundary points. Based on this idea, the approach proposed in [9] allows us to obtain a function on a surface mesh which is harmonic and has 1 maximum, 1 minimum point and, by the Morse–Euler formula, 2\( g \) saddles, with \( g \) the surface genus.

### 3. Reeb graph definition

Reeb graphs are topological constructs that were defined by a mathematician, George Reeb, in his work dated 1946 [10]. Given a manifold \( M \) and a real-valued function \( f \) defined on \( M \), the graph introduced by Reeb, or Reeb graph, is the quotient space defined by the equivalence relation that identifies the points belonging to the same connected component of level sets of \( f \). Under some hypothesis on \( M \) and \( f \), Reeb stated the following theorem, which is actually the definition of what will be called Reeb graph:

\[
\text{Let } M \text{ be a compact manifold of dimension } n \text{ and } f \text{ a simple Morse function defined on } M, \text{ and let us define the equivalence relation } \sim \text{ as } (P, f(P)) \sim (Q, f(Q)) \text{ if and only if } f(P) = f(Q) \text{ and } P, Q \text{ are in the same connected component of } f^{-1}(f(P)).
\]

The quotient space on \( M \times \mathbb{R} \) induced by “\( \sim \)” is a finite and connected simplicial complex\(^2\) \( K \) of dimension 1, such that the counter-image of each vertex \( \Delta^0 \text{ of } K \) is a singular-connected component of the level sets of \( f \), and the counter-image of the interior of each simplex \( \Delta^1 \text{ of } K \) is homeomorphic to the topological product of one connected component of the level sets by \( \mathbb{R} \) [10].

Reeb also demonstrated that the degree of a vertex \( \Delta^0 \text{ of index } 0 \text{ or } n \) is 1 and the index of a vertex \( \Delta^0 \text{ of degree } 1 \text{ is } 0 \text{ or } n \). In case of a surface, this means that maxima and minima, that are vertices of degree 0 and 2 respectively, have index 1, as illustrated in Fig. 5(a) for the case of the maximum. In general, from the theorems in [10], it is possible to derive that leaf nodes of \( K \) can be either maxima or minima of \( f \).

Moreover, if \( n \geq 3 \) the degree of vertices \( \Delta^0 \text{ of index } 1 \text{ or } n - 1 \) is 2 or 3. If \( n = 2 \) the degree of vertices \( \Delta^0 \text{ of index } 1 \text{ is } 2, 3, \text{ or } 4 \). It is also possible to deduce that, for 2-manifolds that can be embedded in \( \mathbb{R}^3 \) (and more in general for orientable surfaces), the degree of vertices of index 1, i.e. vertices representing saddles, is always 3 ([20], page 66), as illustrated in Fig. 5 (b). For non-orientable surfaces, the degree of critical points of index 1 (saddles) may be also 2 or 4.

Finally, the degree of vertices \( \Delta^0 \text{ of index different from } 0, 1, n - 1, \text{ or } n \) is 2. These results clarify the relations between the degree, or order, of the vertices of the simplicial complex \( K \) representing the quotient space and the index of the corresponding critical points.

---

\(^1\) A function defined over polyhedra is called general if \( \forall (u, v), \text{ where } (u, v) \text{ is an edge, } f(u) \neq f(v) \).

\(^2\) We recall that a k-simplex \( \Delta^k \text{ in } \mathbb{R}^n, 0 \leq k \leq n \), is the convex hull of \( k + 1 \) affinely independent points \( A_0, A_1, \ldots, A_k \), called vertices. A finite simplicial complex can be defined as a finite collection of simplexes, together with their faces of any dimension, that can meet only along a common face. where a face of a k-simplex \( \Delta^k \) is a simplex, whose set of vertices is a non-empty subset of the set of vertices of \( \Delta^k \). The dimension of a simplicial complex is the maximum dimension of its simplexes [30].
Fig. 5. Representation of the degree of a vertex corresponding to a critical point of index 0 (a) or 1 (b). The maximum in (a) has degree equal to 1, while the saddle in (b) has degree equal to 3.

Fig. 6. A 2D manifold studied with respect to the height function (left) and the corresponding Reeb graph (right). The letters from a to i denote the correspondence between critical values and Reeb graph nodes.

Table 1
Relations between the genus of a surface $M$ and the number of loops of the corresponding Reeb graph $K$

<table>
<thead>
<tr>
<th>Type of the 2-manifold</th>
<th>Relation between loops and genus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orientable without boundary</td>
<td>$\beta_1(K) = g$</td>
</tr>
<tr>
<td>Orientable with $b_M$ boundary comp.</td>
<td>$g \leq \beta_1(K) \leq 2g + b_M - 1$</td>
</tr>
<tr>
<td>Non-orientable without boundary</td>
<td>$0 \leq \beta_1(K) \leq \frac{g}{2}$</td>
</tr>
<tr>
<td>Non-orientable with $b_M$ boundary comp.</td>
<td>$0 \leq \beta_1(K) \leq g + b_M - 1$</td>
</tr>
</tbody>
</table>

The example in Fig. 6 shows the Reeb graph of a closed surface with respect to the height function. In Fig. 6 (left) some level sets of the height function are drawn. The corresponding Reeb graph is depicted in Fig. 6 (right), where the labels denote the correspondence between the critical values of $f$ and the nodes of the Reeb graph.

For orientable, closed 2-manifolds, the number of loops in the Reeb graph corresponds to the genus of the manifold. This result has been generalized in [31], where the authors demonstrate that the first Betti number of $M$, $\beta_1(M)$, of non-homologous loops of the surface is an upper bound of the number of loops $\beta_1(K)$ of the Reeb graph. The equality holds in case of orientable surfaces without boundary, while, in general, only inequalities are satisfied. Table 1 summarizes the relations between the number loops of the Reeb graph and the genus $g$ of a surface $M$ (2-manifold).

In the case of 3-manifolds, it is not longer true that the number of loops of an orientable, closed 3-manifold is independent of the mapping function $f$. In addition, it has been proven that for every 3-manifold $M$ there exists at least a Morse function $f$ such that the Reeb graph of $M$ with respect to $f$ is a tree [31].
4. Reeb graphs in computer graphics

To the best of our knowledge, Reeb graphs, have been firstly introduced in Computer Graphics by Shinagawa et al. [11]. Notice that in the literature the term Reeb graph is used to identify the simplicial complex associated with the quotient space defined by Reeb.

The Reeb graph has been mainly used for 2-manifolds, which is quite natural in the sense that Computer Graphics is mainly targeted at studying shapes that may be abstracted as manifolds with a natural embedding in $\mathbb{R}^3$. The Reeb graph has been also used for the analysis of 3-manifolds with boundary, and in this case their structure is studied either by introducing a virtual closure of the manifold [32], or by associating a Reeb graph to each 2-manifold boundary component of the 3-manifold and keeping track with a supplementary graph of the changes between interior and void [11,33].

The Reeb graph can be seen as a generalization of the contour tree [34], a popular tool mainly used for spatial data handling and visualization purposes. Contour trees, which are sometimes confused with Reeb graphs in the literature, are rooted in the same background as Reeb graphs, but are presented with less formal and more application-oriented definitions. The relationship between Reeb graphs and contour trees is discussed in Section 4.1.

In the Computer Graphics literature, a number of approaches have been proposed to extract the Reeb graph from polyhedral models. Two main approaches can be identified to solve the typical problem represented by degenerate critical points and non-isolated critical points. A first class of methods delegate the solution of problematic cases to local adjustments or perturbations [11,31,34,35,22]. 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. On the contrary, a second class of methods approach this problem by adopting a semantic characterization of the model and introducing discrete structures like Extended and Multi-resolution Reeb graphs, Level Set Diagrams, etc., which are discussed throughout this section.

4.1. Contour trees

Contour trees are defined for scalar fields $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, where $D$ is a simply-connected sub-domain of $\mathbb{R}^n$. Notice that, in this case, the shape is entirely defined by the function $f$ itself. Moreover, since the domain of definition of $f$ is simply connected, the graph is a tree, as originally shown in [36]. On the contrary, Reeb graphs are defined for a more general class of shapes, $n$-dimensional manifolds, and therefore they can have a general connectivity which reflects the topology of the manifold.

In the literature, several and slightly different definitions of contour trees have been introduced. All of them reflect the intuition that each connected component of the level sets of the scalar field is contracted to a point and the contour tree represents the events in their evolution, as the isovalue spans the range of possible values [34]. These events, which correspond for example to the creation, union, or disappearance of level sets components, correspond to the presence of critical points of the scalar field.

In the following, we report a Morse-theoretic definition of contour tree inspired by the one proposed in [34] because it characterizes well the behaviour of the tree in correspondence to the critical points. An alternative definition can be found in [37]. In particular we adopt the term component-critical points to denote critical points at which only the number of connected components of the level set varies, as used for example in [22]. Given a scalar field $f' = (D, f)$, with $f$ Morse, two level sets $C$ and $C'$ are said to be equivalent if there exists some $f$-monotone path $\alpha$ in $D$ that connects some point in $C$ with another in $C'$ such that no point $x \in \alpha$ belongs to a contour (i.e. a level set component) of any component-critical points of $f$ ([34]). The classes induced by this equivalence are called contour classes. Then, the contour tree is a graph $(V, E)$ such that:

1. $V = \{v_i \mid v_i \text{ is a component-critical point of } f\}$;
2. For each infinite contour class created at a component-critical point $v_i$ and destroyed in another component-critical point $v_j$, $(v_i, v_j) \in E$.

Finally, it is assumed that an arc $(v_i, v_j)$ is directed from the higher to the lower value of $f$ on it. Fig. 7 shows the contour tree of a 2D scalar field.

The differences in the existing definitions of contour trees mainly depend on the type of evolution, that is on the type of critical point, stored in the structure. The contour tree typically keeps track of the critical points in which only
the number of components of the level set varies, but not the genus of level sets. For 2D scalar fields this situation does not occur, but for 3D scalar fields there are critical values at which the topological genus of the isosurface changes without modifying the number of connected components nor the adjacency of the level sets [35,22]. On the contrary, Reeb graphs take into account not only the number but also the changes in the topology of connected components of the level sets. Since the original definition by Reeb states that the counter-image of the interior part of each 1-simplex is homeomorphic to the topological product of one connected component of the level sets and \( \mathbb{R} \), no changes in the topology of the level sets are admitted along an edge of the graph. To give an example, let us consider a scalar field in \( \mathbb{R}^3 \) whose isosurface changes genus across a critical value of the scalar field. In this case, the plain contour tree would always count one connected component while the Reeb graph would identify one critical point having index 2 and therefore reflects its presence in the related graph structure. In other words, the contour tree does not code all the saddles that might be identified by analyzing the level set evolution [35,22].

It is worth noticing however that some authors have proposed to enrich the contour tree with further information on all topological changes of the level sets, by adding a node also in correspondence to critical values where not the number but the topology of the level sets changes. This tree was firstly introduced in [35] with the name of Augmented Contour Tree (ACT) but it was renamed contour topology tree by [22] to avoid confusion with the augmented contour tree described in [38], which denotes the subdivision of the contour tree by all vertices of the input mesh.

4.2. Extended and discrete Reeb graph

Firstly introduced in [39] for terrain models, the Extended Reeb Graph (ERG) is able to represent a surface through a finite set of level sets of a given mapping function \( f \). The definition of the Reeb graph has been extended in [40] to triangle meshes representing 2-manifolds embedded in \( \mathbb{R}^3 \), with or without boundary. An important property of ERG is that it can also deal with vertices that corresponds to degenerate configurations of critical points, such as volcano rims [41].

More formally, we denote by \([f_m, f_M]\) an interval containing the co-domain of a continuous \( f \) defined on the surface \( M \) such that \( f_m = f_0 < f_1 < \cdots < f_{np} < f_M = f_{np+1}, np \in \mathbb{N}, \) and \( f^{-1}(f_i), i = 0, \ldots, np + 1 \) are all non-degenerate level sets. We consider

\[
I = \{(f_i, f_{i+1}), i = 1, \ldots, np\} \cup \{f_0, f_1, \ldots, f_{np}, f_{np+1}\}
\]

a partition of the interval \([f_m, f_M]\) provided by the set of the \( np + 1 \) open interior parts of \([f_m, \ldots, f_M]\) and the function values of the isocontours \( \{f_i\} \). Then, an extended Reeb equivalence between two points \( P, Q \in M \) is defined by:

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

Therefore, it follows that all points belonging to the counter-image of an element \( t \in I \) 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 a discrete space, called Extended Reeb (ER) quotient space. As described in [40] the ER quotient space, which is an abstract sub-space of \( M \) and is independent of the geometry, may be represented as a traditional graph named the Extended Reeb Graph (ERG), by following the adjacency relations among the elements of \( f^{-1}(f(t)), t \in I \).
Fig. 8. Pipeline of the ERG extraction. (a) Evaluation of the values of a function $f$ on a surface model, from blue values representing minima to red values representing maxima; (b) the recognition of critical areas: minima (blue), maxima (red) and saddles (green); (c) the expansion process; (d) the resulting ERG configuration. This model comes from the AIM@SHAPE repository: http://shapes.aim-at-shape.net/. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Beside the combinatorial representation of the graph, the ERG may be embedded in $\mathbb{R}^3$ by associating to each node the position of the barycenter of the corresponding region and visualized almost everywhere as an approximation of the centerline skeleton of $M$.

In [42], both the definition and the algorithm have been extended to surfaces having an arbitrary number of boundary components with the aim of obtaining a minimal (in the sense of the number of graph loops) representation of the ERG. This method departs from the consideration that the genus of a surface with boundary is traditionally defined as the genus of the closed surface obtained closing every boundary component with a disk [43] and virtually closes all boundary components. Thus the proposal in [42] differs from [31], where no closures of the boundary components are introduced.

From an algorithmic point of view, the ERG extraction is based on a generalized surface characterization aimed at a region-oriented rather than a point-oriented classification of the behaviour of the surface [44]. Since the pre-images of the values $f_i, i = 1, \ldots, np$ decompose $M$ into a set of regions, each region is defined a regular or a critical area according to the number and the value of $f$ along its boundary components. Critical areas are classified as maximum, minimum and saddle areas and correspond to nodes of the graph. Then arcs between nodes are detected through an expansion process of the critical areas, which tracks the evolution of the isocontours. The pipeline of the ERG extraction is illustrated in Fig. 8.

The relationship between the genus of the mesh and the cycles in the ERG is maintained, as discussed in [40,42], even in case of rough samplings of the co-domain of $f$. In fact, during the graph extraction, the algorithm automatically checks that all regions on $M$ induced by the level sets $f_i$ do not contain holes, by counting them by means of the Euler formula $\#\text{vertices} - \#\text{edges} + \#\text{faces} = 2(1 - \#\text{holes}) - \#\text{boundaries}$, where $\#\text{boundaries}$ represent the number of boundary components of a region. In case a region contains a hole an additional level set is locally considered.

On the basis of the ERG representation, a further extension of the domain of the Reeb graph to unorganized point clouds of 3D scan data has been proposed in [45]. The assumption on the point clouds is that they represent a human body. The limitation that the original data are not organized in a polygonal mesh is overcome assuming that the Euclidean distance among a point $p$ and its closest point $q$, is smaller than a given threshold $\epsilon$, $d(p, q) < \epsilon$. Point sets whose sampling is sufficiently fine are connected in a discrete sense. Therefore, level sets are defined as points that share the same value of the mapping function and are connected in the discrete sense. The resulting graph is called the Discrete Reeb Graph (DRG).

4.3. Multi-resolution and augmented Reeb graph

Hilaga et al. in [46] provide a Multi-resolution Reeb Graph (MRG) representation of triangle meshes which is independent of the object topology. In fact, the MRG is constructed through a multi-resolution slicing strategy that
Fig. 9. Pipeline of the MRG extraction. (a) Subdivision of the domain of the mapping function $f$ (the height function in this example) into a set of intervals; (b) subdivision of triangles so that the image of every triangle belongs to only one interval; (c) creation of a node for each triangle set, that is the set of connected components of triangles whose images belong to the same interval; (d) detection of arcs by checking the region adjacency of triangle sets.

takes advantage of the observation that, when halving the distance between two level sets, the nodes of the graph keep a hierarchical relationship. The power of this approach is clear: the surface shape can be processed at different levels of detail and an estimation of the object features is automatically provided. An innovative mapping function that depends on the average of the geodesic distance of a point from the other vertices is also defined. The construction of the graph begins with the extraction of the graph at the finest resolution desired, then adjacency rules are used to complete the multi-resolution representation in a fine-to-coarse order. Notice that, in this setting, no topological controls are performed during the graph extraction; therefore, there may be no correspondence between the number of loops of the MRG and the number of tunnels of the model. The algorithm for extracting the MRG is shown in Fig. 9.

The approach proposed in [47] works on a volumetric model. In this case, the data are swept with a set of parallel planes generating a set of slices, which are formed by the sets of grid elements bounded by two adjacent isosurfaces. Each connected component of a slice is called a ribbon while the contours are given by the intersection of the isosurfaces with a set of slicing planes. Both ribbons and contours correspond to nodes of the Reeb graph while their adjacency is coded in the edges. The graph described in this approach is called an Augmented Reeb Graph (ARG) because it codes also geometry information for each contour and each ribbon. To avoid that an object handle is completely contained within a ribbon, the Euler characteristic of each isosurface component is computed and, eventually, the sweep is locally refined. In this way the topology of the volume is completely coded and, in each interval, there is the correspondence of the Reeb graph structure with its Euler characteristic.

4.4. Level set diagrams

In the Level Set Diagrams (LSD) of triangulated surfaces proposed in [48], each contour is visualized through its centroid. The construction of the LSD uses Euclidean distances for wave propagation from a seed point. To automatically select the source point an heuristic is used that determines a privileged “slicing direction”. In this approach, a skeleton-like structure, available for input objects of genus zero, is proposed, which is essentially a tree made of the average points associated with the connected components of the level sets of the geodesic distance from the source. An extension of the approach in [48] to non-zero genus surfaces has been presented in [49]. Since a source point is chosen for propagating contours on the shape, LSDs perform as well as the shape is tubular and the sections are circular. However, this approach is potentially useful for medical applications such as virtual colonoscopy, visualization of blood vessels, computer-driven angiography, in order to visualize networks of blood vessels or branching patterns of air passageways in the lungs. An example of the Level Set Diagram for a vessel model is shown in Fig. 10.

The method in [50] is based on the use of a multi-scale curvature evaluation for locating seed points, which is achieved by using a set of intersection curves between the input surface and a collection of spheres centered in the mesh vertices and with increasing radius. Then, seed points are sequentially linked by using a wavefront traversal distance defined for a simplicial complex. In particular, we highlight that the set of seed points corresponds to the vertices of the graph that has degree 1. Therefore, the complexity of the graph depends on the seed points and the number of curvature scales chosen. With reference to the example shown in Fig. 11, all extrema of the cactus are recognized as high-curvature points. This depends on the fact that, when a large-curvature scale is used, also flat points like those on the basis of the cactus may be considered as highly curved, depending on the ratio of the radius of the sphere to the length of its intersection with the model. On the contrary, the hand shown in [51] has a large flat region which is never recognized as a high-curvature point, thus leading to a different graph representation.
4.5. Computational aspects

Several techniques have been reported in the literature for the computation of Reeb graphs and related structures. Despite its general definition, existing algorithms for Reeb graph extraction mainly work on 2D and 3D manifolds, and only recently 3D time-dependent data [32]. In this section, we overview the algorithm proposed by Shinagawa and Kunii [52] and the computationally optimal method proposed in [31]. Specific algorithms for different structures (such as contour trees, ERG, MRG, ARG, LSD, etc.) are provided in the corresponding papers and their technical aspects are discussed and summarized in a different survey contribution [19]. A summary of the computational costs of the methods mentioned in this paper is reported in Table 2.

The first algorithm to compute Reeb graphs, proposed by Shinagawa et al. [52], automatically constructs the graph from surface contours generated by the height function. A weight function, which depends on the average distance between the vertices of two different contours, is defined for each pair of contours lying on adjacent (consecutive) level sets and coupled with the a priori knowledge of the surface genus to determine connections.

A general algorithm for the Reeb graph extraction of 2-manifolds with or without boundary represented by a simplicial complex has been proposed in [31]. This approach also works for non-orientable models, like the Klein bottle. The basic assumption here is that the mapping function is Morse and simple, so that critical points have pairwise different function values. Then, the Reeb graph is constructed by storing the level sets while sweeping the domain of the function. Critical points are identified analyzing the star of each vertex and classified according to the approach in [14]. Once all critical points have been detected, all vertices of the model are processed according to the increasing value of the function \( f \) and the evolution of level sets is tracked. Since operations are done on the edges, the complexity of the algorithm is \( O(n \log n) \), where \( n \) is the number of edges of the complex.
Table 2
Methods for Reeb graphs and related structure extraction

<table>
<thead>
<tr>
<th>Approach</th>
<th>Domain</th>
<th>Output</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>[11]</td>
<td>X</td>
<td>Reeb graph</td>
<td>$O(n^2)$</td>
</tr>
<tr>
<td>[31]</td>
<td>X</td>
<td>Reeb graph</td>
<td>$O(n \log(n))$</td>
</tr>
<tr>
<td>[53]</td>
<td>X</td>
<td>Contour tree</td>
<td>–</td>
</tr>
<tr>
<td>[54]a</td>
<td>X</td>
<td>Contour tree</td>
<td>$O(N \log N)/O(N^2)$</td>
</tr>
<tr>
<td>[37]</td>
<td>X</td>
<td>Contour tree</td>
<td>$O(N \log(N))$</td>
</tr>
<tr>
<td>[22]</td>
<td>X</td>
<td>Contour tree</td>
<td>$O(N + c \log(c))$</td>
</tr>
<tr>
<td>[38]b</td>
<td>X</td>
<td>Contour tree</td>
<td>$O(N \log(N))$</td>
</tr>
<tr>
<td>[34]b</td>
<td>X</td>
<td>Contour tree</td>
<td>$O(C \log(C) + N \log(n))$</td>
</tr>
<tr>
<td>[35]c</td>
<td>X</td>
<td>ACT</td>
<td>$O(n + c \log(n))$</td>
</tr>
<tr>
<td>[44,40]</td>
<td>X</td>
<td>ERG</td>
<td>$O(\max(m + n, n \log(n)))$</td>
</tr>
<tr>
<td>[39]</td>
<td>X</td>
<td>ERG</td>
<td>$O(\log(n))$</td>
</tr>
<tr>
<td>[45]</td>
<td>Point clouds</td>
<td>DRG</td>
<td>$O(n)$</td>
</tr>
<tr>
<td>[46,55]</td>
<td>X</td>
<td>MRG</td>
<td>$O(n + m)$</td>
</tr>
<tr>
<td>[47]</td>
<td>X</td>
<td>ARG</td>
<td>$O(n \log(n))$</td>
</tr>
<tr>
<td>[48,49]</td>
<td>X</td>
<td>LSD</td>
<td>$O(n \log(n))$</td>
</tr>
<tr>
<td>[50]d</td>
<td>X</td>
<td>Skel. Centerline</td>
<td>$O(n)$</td>
</tr>
<tr>
<td>[56]</td>
<td>X</td>
<td>Reeb graph</td>
<td>$O(N) + O(nc) + O(c^2)$</td>
</tr>
<tr>
<td>[57]</td>
<td>X</td>
<td>Volume skeleton</td>
<td>$O(N) + O(nc) + O(c^2)$</td>
</tr>
</tbody>
</table>

Symbols: $n$ represent the number of vertices or points; $N$ the number of simplexes; $C$ the number of tree nodes; $\alpha$ is the inverse of the Ackermann function; $c$ the number of critical points; $m$ the number of vertices inserted in the mesh during a possible contouring phase.

- The complexity is, respectively, $O(N \log N)$ for 2D and $O(N^2)$ for 3D domains.
- The method is essentially the same in both contributions. The cost improvement is due to a more efficient implementation of the data structures.
- The complexity of this method depends on the use of an FIFO or a priority queue.
- Once a set of seed points have been recognized, the complexity of the skeleton extraction is linear in the number of mesh vertices (because it is based on a region-growing algorithm where each triangle of the mesh is visited once), but an accurate evaluation of the high-curvature points has quadratic cost [58].

5. Applications of Reeb graphs

One of the most important features of Reeb graphs is that they provide a structure that effectively codes shapes from both a topological and a geometrical perspective. Topology here means that the shape can be described as a configuration of parts that are attached together respecting the topology of the shape, while geometry means that the different parts correspond to features of the shape, as embedded into the Euclidean space, that have specific properties and descriptive power (e.g., protrusions, elongated parts, wells). As a consequence of their ability to extract high-level features from shapes since their introduction in Computer Graphics, Reeb graphs have been gaining popularity as an effective tool for different shape analysis and description tasks.

Application fields related to the use of Reeb graphs are surface analysis and understanding [11,44]; identification of topological quadrangulations [49]; data simplification [39]; animation [59]; human body segmentation [45]; surface parameterization [60,61] (Fig. 12); object reconstruction [62] and editing where the stored information is exploited for recovering. Moreover, the knowledge of the shape topology given by the graph structure improves the tiling from contour lines [62], thus solving the correspondence and the branching problems.

The compactness of the 1D structure, the natural link between the function and the shape, and the possibility of adopting different functions for describing different aspect of shapes have led to a massive use of Reeb graphs for similarity evaluation, shape matching and retrieval [46,63]. In [46] the Multi-resolution Reeb Graph is used in a multi-resolution fashion for shape matching. The ratios of the area to the length of the model sub-part in the whole model are associated with each node of the graph, and are used as attributes during the graph-matching phase. The set of the geometric attributes is further enriched in [55], where, for each slice, the authors consider the volume, a
Fig. 12. Surface parameterization using the Extended Reeb Graph. ((a), (b)) A topology-based decomposition of the shape derived from the ERG is used to define a chart decomposition of the mesh, and each chart is parameterized with respect to the cuts shown in (b); (c) the normal-map images.

Fig. 13. (b) Sub-part correspondence detection between the two mechanical parts shown in (a) with their ERG structures superimposed.

statistical measure of the extent and the orientation of the triangles, an estimation of the Koenderink shape index and a statistic of the texture.

The application of the Extended Reeb Graph to database retrieval has been addressed in [63] and discussed in the context of CAD models in [64]. The decomposition into significant regions induced by the ERG defines a structural description of the shape, which is coupled with an error-correcting subgraph isomorphism to build up a shape retrieval system. In particular, the proposed graph-matching framework makes it possible to plug in heuristics for tuning the algorithm to the specific application and for achieving different approximations to the optimal solution. In [64] the authors discuss a method for measuring the similarity and recognizing sub-part correspondences of 3D shapes, based on the coupling of a structural descriptor, like the ERG, with a geometric descriptor, like spherical harmonics. The proposed framework offers an effective solution to the problem of searching and retrieval in general, and, since the matching relies on structural descriptions, the system is particularly suitable for sub-part correspondence (Fig. 13). It is discussed how a structure-based matching can improve the retrieval performance in terms of both functionalities supported (i.e., partial and sub-part correspondences) and the variety of shape descriptions that can be used to tune the retrieval with respect to the context of application.
A comparison between the performance of the Extended Reeb Graph and Multi-resolution Reeb Graph on a database of 280 3D models is presented in [65], in the context of 3D shape classification. Using the same database, we propose here a comparison between the retrieval performance of ERG and MRG, in terms of the precision-recall diagram [66], see Fig. 14.

Since Reeb graphs generalize to $n$-dimensional manifolds the concepts behind contour trees, their application domains partially overlap, for instance scientific visualization. The original application of contour trees was in topography [53] and Geographic Information Systems (GIS) [56]. As far as 3D scalar fields are concerned, main applications of contour trees are automatic isosurface propagation and scientific visualization [35,57,34]. In this context, to overcome the problem of the size of the graph, the work in [67] proposes a method for extracting a multi-resolution contour tree and visualizing the tree sub-parts (the so-called branches). Branches, that are chains of connected arcs, are sorted according to their importance into the contour tree and progressively rendered into 3D space moving each node to a $z$ value that corresponds to the value of the scalar field in that node. More recently, these branches have been used to topologically simplify the rendering of complex volumetric datasets [68,69]. Finally, contour trees in higher dimensions apply to X-ray analysis and visualization [38] and to scientific visualization [70,71].

6. Conclusions and emerging trends

As observed before, Reeb graphs play a fundamental role in the field of computational topology for shape analysis. The idea of exploring the topology of a space by analyzing the behaviour of a possibly varying real function defined on it finds its root in the classical Morse theory, and it is common to other popular descriptors in computational topology, namely Morse and Morse–Smale complexes [9] and tools in size theory [72] and persistent homology theory [8]. The reader is referred to [19] for a detailed analysis of the differences in the shape description and interpretation offered by these methods.

The distinctive property of Reeb graphs is that they encode topological information on shapes in a 1D structure, disregarding the dimension of the manifold representing the shape. As far as the study of the shape of orientable manifolds up to dimension 3 is concerned, the computation and use of Reeb graphs are quite at a mature stage. However, there is an increasing availability of large multi-dimensional data sets, both static and dynamic, arising for example from scientific simulations. Therefore, one of the main challenges is to compute Reeb graphs for spaces of dimension higher than 3, in order to be able to extract knowledge from high-dimensional data. The problem of analyzing the variation of the shape description for dynamical shapes or phenomena deserves special mention: from a theoretical point of view, the dependency on the time could be handled as any another variable, but special care has to be taken for the interpretation of dynamical effects. In this context, there are some works that have already proposed algorithms for the computation of time-varying contour trees and Reeb graphs [73,32,71].
Regarding the effectiveness of Reeb graphs for shape description and comparison, we have seen that they provide a general framework for studying a shape which is parameterized with respect to the mapping function used. The mapping function plays the role of the *lens* through which we look at the properties of the shape, and different functions provide different insights, fixing properties that are recognized by the Reeb graph [51]. A very interesting problem is how to devise mapping functions suites that can be formalized beyond a generic best practice or rule of thumb. These functions should be able to capture features that are somehow orthogonal not only in an abstract conceptual space but also in a geometric sense. In this field there is also a growing interest in defining measures that are able to determine how much and where two or more functions defined on the same model differ [74,75]. We think that one of the most promising research direction is the use of Laplacian eigenfunctions, that are intrinsic to the shape and encapsulate the idea of orthogonality of shape description [76]. The first proposal involving the use of the Laplacian operator for the computation of Reeb graphs is the one of Steiner & Fischer [77], who extract a Reeb graph representation from a set of contours obtained by solving a mesh Laplacian system. Related concepts are also dealt with in [78], where a Morse function with a user-controlled number and configuration of critical points is computed by solving a relaxed form of Laplace’s equation, and is used to build Morse complexes. Our idea is to interpret the eigenfunctions of the Laplacian operator computed on a mesh as real-valued functions defined on the mesh itself, and use them as mapping functions to extract a collection of Reeb graphs. This collection could be used in turn for shape comparison and retrieval purposes. Although we believe in the potential of Laplacian eigenfunctions as shape descriptors, there is a number of questions that should be addressed in order to use them for shape retrieval purposes, such as the problem of eigenfunction *switching* [79]. This is an actual and open research topic.

A related interesting issue concerns the way to concurrently use multiple functions. In this scenario, we foresee the definition of new descriptions that extend Reeb graphs to multi-skeletons or even combined skeletons, possibly augmented with sufficient geometric data which would capture the most relevant features and allow an efficient processing of the shape data.

**Acknowledgements**

The authors thank L. De Floriani, P. Frosini, C. Landi and L. Papaleo for fruitful discussions. The authors also thank the anonymous reviewers for their helpful comments and suggestions.

This work has been developed in the CNR research activity (ICT-P03) and partially supported by the European Network of Excellence FP6 IST NoE 506766 “AIM@SHAPE”.

**References**


