Topology Based Selection and Curation of Level Sets

Chandrajit Bajaj, Andrew Gillette, and Samrat Goswami

Center for Computational Visualization, University of Texas at Austin Austin, Texas 78712
bajaj@cs.utexas.edu, agillette@math.utexas.edu, samrat@ices.utexas.edu

Summary. The selection of appropriate level sets for the quantitative visualization of three dimensional imaging or simulation data is a problem that is both fundamental and essential. The selected level set needs to satisfy several topological and geometric constraints to be useful for subsequent quantitative processing and visualization. For an initial selection of an isosurface, guided by contour tree data structures, we detect the topological features by computing stable and unstable manifolds of the critical points of the distance function induced by the isosurface. We further enhance the description of these features by associating geometric attributes with them. We then rank the attributed features and provide a handle to them for curation of the topological anomalies.

1 Introduction

Problem and Motivation.

The selection of isosurfaces (or alternatively level sets of a trivariate function) for the visualization of volumetric imaging or simulation data, is often subject to the constraints imposed by the application domain. While some applications put more stringent criteria on this process than others, every application requires an analysis of topological and geometrical information to ensure that an appropriate choice is made. Topological constraints require that the extracted level set should have a certain topology. Naturally, this problem has drawn the attention of researchers for a long time. As a result, the basic topological information of level sets can be unambiguously encoded by the well-known data structure, called the Contour Tree [31]. The Contour Tree (CT) is a topological description of the entire volumetric data, and furthermore encodes the information of the number of connected components.

While extremely useful, CT does not capture all the necessary information about the topology of the level sets. In particular, information about the combinatorics and topology of tunnels (complementary space) of every connected component of the level set is not encoded in CT. To overcome this deficiency, an enhanced CT data structure was proposed by Pascucci and Cole-McLaughlin [29], who gave an

algorithm to further detect and store some additional information of the level set, in an Augmented Contour Tree (ACT).

The ACT, while appropriate in encoding the topology related to the first 3 Betti numbers [29] of each level set, it does not give a quantitative measure of the level set and its complementary space features, to aid in its selection or curation . This is the main motivation of our work. In this paper, we will show that suitable selection of an isosurface in any application context must be additionally guided by a close examination of the complementary space to the primal domain. Further, we will show how relevant topological and geometrical features of a selected isosurface can be calculated, ordered, visualized, and, if necessary, curated.

The significance of such complementary space features can be seen in the isosurface choices for the simple example of the ion channel Gramicidin A, obtained from soil bacteria Bacillus brevis. The initial molecular surface selection process makes use of the CT. In Figure 1 (a.1-4) we show two isosurfaces extracted from the same edge of the contour tree for the volumetric data which has been synthetically created from the atomic model obtained from Protein Data Bank (PDB) [5]. The PDB entry of this molecule is 1MAG. Figure 1 (a.1) shows a snapshot of the volume rendering of the data in our in-house interrogative volume visualization software VolRover [13] along with the CT shown in the bottom panel. In Figure 1 (a.2) we show the molecular surface selected using CT which does not accommodate any tunnel through the molecular surface. Figures 1 (a.3,4) show two views of another molecular surface selected from the same scalar volume which, on the other hand, has the ion channel through the molecule present. This means the choice of the first isosurface is incorrect. Note that there is no change in the number of components as one contour is deformed into the other, hence the CT is not sufficient in guiding our choice of isovalue. Further as can be seen in Figure 1 (a.4), the surface accommodates two more small tunnels (marked by black circles) which are merely the artifacts of the selection process. Using the algorithm presented in this paper, we can eliminate those small tunnels.

Figure 1 (b.1-3) shows another example of the isosurface of the molecule mouse Acetylcholinesterase (mAchE). In this case, the active site of the molecule is buried deep inside a depression on the molecular surface. We will later computationally define such depressions and we will call them "pockets". Correct extraction of the molecular surface of mAchE, in this case, requires that the pocket near the active site is properly preserved and also it should not be too narrow, e.g. narrower than the diameter of Acetylcholine molecule which has to pass through the opening of the pocket in order to bind to the active site of mAchE. Figure 1 (b.1) shows all such pockets identified on the molecular surface of mAchE. The small pockets on the surface arising from incorrect selection of isovalue, are not desired because they adversely affect the calculation of molecular energetics. Using the algorithm presented in this paper, we can rank the pockets based on their geometric attributes (volume, for example) and that helps deleting those small pockets. Figure 1 (b.2) shows the surface after selective removal of all the small pockets keeping only three significant ones which include the one near the active site. Closeup of the main pocket is shown in Figure 1 (b.3).

The selection of an isosurface should also aim to preserve inherent symmetry in the data. In Figure 1 (c.1-3) we show two isosurfaces for the Nodavirus. This icosahedral virus infects the central nervous system of fish and causes a disease called viral nervous necrosis. The main point of contention while selecting an isosurface

is preserving the symmetry of the viral capsid. We show the volume rendering of the cryo-EM density map of the virus in Figure 1 (c.1). Figure 1 (c.2) shows one plausible isosurface that does not respect the symmetry. In particular, near the points of five-fold rotational symmetry, we ought to have five tunnels arranged in a circle or none at all. In Figure 1 (c.3) we show another isosurface which actually reflects the symmetry properly. Without careful investigation of the complementary space topology, aided by the algorithm described in this paper, it is not possible to apply the knowledge of these features to guide the isosurface selection. Therefore, we seek both topological and geometrical details of the complementary space to an extracted isosurface.

The overall isosurface extraction process from a scalar volume is thus guided by the knowledge of both topology and domain-specific geometry of both the primal and complementary space. As in [19], we also consider the homology as a measure of topological complexity of the extracted level set. We further consider the depressions on the surface as added complexity as they play a bio-chemically vital role in the context of selection of molecular interface [18]. We call these depressions *pockets*. The main contribution of this work is the attachment of geometry and possibly other domain-specific attributes with the detected topological features that guides the subsequent curation process.

Prior work.

Systematic interrogation of topological and geometric attributes of the level sets over a range of isovalues started with the introduction of two powerful data structures Contour Spectrum (CS) [4] and Contour Tree (CT) [31]. While CS focuses mostly on the differential and integral attributes like, area, volume, curvature etc., CT encodes topological properties like, number of connected components etc., of the contours over a range of isovalues. Historically, CT was first introduced in [6]. It was first used with regard to isosurfaces by van Kreveld et al. [31]. They also used CT to compute seed sets, which help generate isosurfaces efficiently. Further work by Carr et al. proved that contour trees of any dimension could be computed in $O(n \log n)$ time, where n is the number of simplices in the geometrical decomposition [8]. Pascucci and Cole-McLaughlin expanded the topological information in the CT data structure by adding Betti number information to each edge of the graph [29]. Recently, CT has been further enhanced by adding information about domain specific geometric attributes and such multi-attributed CT (MACT) proved to be very useful in analyzing the bio-chemical properties of the molecular interfaces [35]. Carr, Snoeyink and van de Panne simplified the CT based on local geometric data [9]. The algorithm in [9], although different from our approach, can be used to further facilitate in simplification of the volume and selection of the isosurface.

Due to the fundamental nature of the problem, topology simplification of geometric models have received attention from outside of visualization community as well. In order to achieve controlled topological simplification of triangulated geometry, Guskov and Wood [25] grow a small ball of radius ϵ on the surface to detect small tunnels and remove them by cutting the mesh. Wood et al. [32] use the Reeb graph of a *height* function to detect and delete small handles from an isosurface. El-Sana and Varshney have worked on topology controlled simplification of CAD models where they first detect the crease edges and roll a ball of small radius to identify the holes which do not allow the ball to pass through [20]. Nooruddin and

Turk have proposed an algorithm that converts a model into volumetric data and apply dilation and erosion to perform simplification [28]. It is to be noted, that all these techniques fall short of depicting the symmetry and a proper ranking of the geometry of the depressions and tunnels of the geometry to be simplified.

A fundamentally different approach was due to Edelsbrunner et al. [19] who proposed the notion of *persistence* in the context of the alpha complex to detect topological features which prevail while the alpha complex undergoes a filtration. A series of results followed along the same line which formalized the notion of persistent homology in order to distinguish between stable topological features from unstable or noisy features [7, 12, 37, 36]. A similar notion of persistence has also proved to be useful in detecting short-lasting and noisy topological features in the context of witness complex [15]. At this stage it is important to note the novelty of our approach. We compute the topological features related to the homology group of the level set and attach geometric attributes which are often meaningful in the context of application the scalar volume has originated from.

The key ingredient of our algorithm in ranking the topological features of the extracted level set is the *distance function* over $\mathbb{R}^3$. The distance function has been used earlier for reconstruction and image feature identification [1, 10, 17, 21]. Chazal and Lieutier [11] have used it for stable medial axis construction. Dey, Giesen and Goswami have used distance function for object segmentation and matching [16]. Goswami, Dey and Bajaj have used it for detailed feature analysis of shape via an annotation of flat and tubular features in addition to shape segmentation [22]. Recently, Bajaj and Goswami have shown a novel use of distance function, induced by a molecular surface, in order to detect secondary structural motifs of a protein molecule [2]. The close connection between the critical point structure of the distance function and the topology of the surface, and its complement, is what we utilize to detect and remove small topological artifacts.

Approach.

The main contribution of this work is the systematic use of the distance function induced by an isosurface, to geometrically complement the encoding of the topology by the Contour Tree, in yielding a curated, selection. With our new approach the selected isosurface is extracted, and then filtered, with the aid of the critical point structure of the distance function, which allows detection and a geometrical ranking of the complementary structure of the isosurface, i.e. the *tunnels* and *pockets*.

First, a suitable isovalue is selected using CT in order to select an isosurface with the required number of components. In the case of molecular interface selection, the number of components is always one. The subsequent computations based on the distance function are then applied to detect the tunnels and pockets. Finally, these features are ranked according to some domain-aware "importance" function which usually quantifies the geometric attributes of those features, and thereby allows the removal of insignificant ones.

We first give a brief description of the distance function, here. Given a compact surface Σ smoothly embedded in $\mathbb{R}^3$, a distance function h_Σ can be designed over $\mathbb{R}^3$ that assigns to each point its distance to Σ.

$$h_\Sigma \ : \ \mathbb{R}^3 \to \mathbb{R}, \ \ x \mapsto \inf_{p \in \Sigma} \|x - p\|$$

In this context, Σ is the level set. For the ease of computation, we approximate h_Σ by h_P which assigns to every point in $\mathbb{R}^3$, the distance to the nearest point from the set P which finitely samples Σ.

$$h_P \ : \ \mathbb{R}^3 \to \mathbb{R}, \ \ x \mapsto \min_{p \in P} \|x - p\|$$

We identify the maxima and index 2 saddle points of h_P which lie outside the level set. The stable manifolds of these critical points help detect the tunnels and the pockets of Σ. Additionally these stable manifolds are used to compute geometric attributes of the detected topological features that they correspond to. Thus we obtain a description of the isosurface, and its complement, in terms of its topological features quantified by their geometric properties, based on which the insignificant features are removed.

2 Preliminary

2.1 Contour Tree

Isosurfaces and contour trees are derived from scalar fields. A scalar field can be characterized as a domain M and a function $f : M \to \mathbb{R}^1$. In differential topology and Morse theory, the critical values of f are formally defined as those values $r \in \mathbb{R}^1$ for which the derivative map df_x is not surjective for some point $x \in f^{-1}(r)$ (see [24], for example) Put differently, r is a critical value of f if and only if $f^{-1}(r)$ is not a manifold of dimension $\dim(\mathrm{M}) - 1$. Each level set $f^{-1}(r)$ is a collection of contours and the topology of these contours allows us to create the contour tree. We note that in all of our examples, our domain M will be $\mathbb{R}^3$, ensuring that our data structure is in fact a tree and not the more general Reeb graph.

The unqualified term "contour tree" refers to a data structure created from a subset of the critical values of f. The first such data structure to be computed efficiently was a "minimal" contour tree by de Berg and van Kreveld [14]. A minimal contour tree has nodes only for isovalues at which contours emerge, split, merge, or vanish. The edges of a minimal contour tree connect nodes along which a contour smoothly deforms and hence indicate the evolution of a contour over a range of isovalues. A minimum contour tree of any dimension can be computed in $O(n \log n)$ time as was proved by Carr et al. in [8]. Such trees can be used to compute seed sets, that is, a set of points from which all contours of a particular level set can be generated [31].

To capture more topological information, the augmented contour tree, as defined by Carr, Snoeyink and Axen, was introduced in [8]. In their terminology, the augmented contour tree refers to a contour tree with nodes for all values in the range of the scalar field, not just the critical values. Thus, the augmented or "full" contour tree can be reduced to the minimal contour tree by removing all degree two nodes. Pascucci and Cole-McLaughlin expanded the data structure by attaching the Betti numbers of each contour in a level set to its corresponding edge in the "full" contour tree [29]. The Betti numbers of a surface, however, are a strictly topological feature and thus do not indicate the geometrical significance of the tunnels and voids that they count. Moreover, it is not clear how to use this data structure in order to selectively remove some undesired topological artifacts.

2.2 Voronoi-Delaunay

In this paper we always assume the distance metric to be Euclidean unless otherwise stated. For a finite set of points P in $\mathbb{R}^3$, the Voronoi cell of $p \in P$ is

$$V_p = \{x \in \mathbb{R}^3 \,:\, \forall q \in P - \{p\},\, \|x - p\| \leq \|x - q\|)\}.$$

If the points are in general position, two Voronoi cells with non-empty intersection meet along a planar, convex Voronoi facet, three Voronoi cells with non-empty intersection meet along a common Voronoi edge and four Voronoi cells with non-empty intersection meet at a Voronoi vertex. A cell decomposition consisting of the *Voronoi objects*, that is, Voronoi cells, facets, edges and vertices is the Voronoi diagram $\operatorname{Vor} P$ of the point set P.

The dual of $\operatorname{Vor} P$ is the Delaunay diagram $\operatorname{Del} P$ of P which is a simplicial complex when the points are in general position. The tetrahedra are dual to the Voronoi vertices, the triangles are dual to the Voronoi edges, the edges are dual to the Voronoi facets and the vertices (sample points from P) are dual to the Voronoi cells. We also refer to the Delaunay simplices as *Delaunay objects*.

2.3 Critical Points of h_P

The distance function h_P induces a flow at every point $x \in \mathbb{R}^3$. This flow has been characterized earlier [21, 22]. See also [17]. For completeness we briefly mention it here.

The critical points of h_P are the points in $\mathbb{R}^3$ which lie within the convex hull of its closest points from P. It was shown by Siersma [30] that the critical points of h_P are the intersection points of the Voronoi objects with their dual Delaunay objects (Figure 2).

- *Maxima* are the Voronoi vertices contained in their dual tetrahedra,
- *Index 2 saddles* lie at the intersection of Voronoi edges with their dual Delaunay triangles,
- *Index 1 saddles* lie at the intersection of Voronoi facets with their dual Delaunay edges, and
- *Minima* are the sample points themselves as they are always contained in their Voronoi cells.

In this discrete setting, the index of a critical point is the dimension of the lowest dimensional Delaunay simplex that contains the critical point.

At every $x \in \mathbb{R}^3$, a unit vector can be assigned that is oriented in the direction of the steepest ascent of h_P. The critical points are assigned zero vectors. This vector field, which may not be continuous, nevertheless induces a flow in $\mathbb{R}^3$. This flow tells how a point x moves in $\mathbb{R}^3$ along the steepest ascent of h_P and the corresponding path is called the *orbit* of x.

For a critical point c its stable manifold is the set of points whose orbits end at c. The stable manifold of a maximum is a three dimensional polytope whose boundary is composed of the stable manifolds of the index 2 saddle points which in turn are bounded by the stable manifolds of index 1 saddle points and minima. See [16, 21] for the detailed discussion on the structure and computation of the stable manifolds of the critical points of h_P.

2.4 Betti Numbers

The i-th Betti number of a manifold is formally defined as the rank of its i-th homology group, H_i. Homology groups are quotient groups; H_i is the i-th cycle group modulo the i-th boundary group. Therefore, H_i is the free abelian group generated by cycles of i-chains that are not boundaries of $(i+1)$-chains. Hence, the i-th Betti number counts the number of independent (i.e. non-homologous) non-bounding cycles. Based on these definitions, we have the following informal notions of Betti numbers for 2-manifolds. The 0-th Betti number equals the number of connected components, the 1-st Betti number equals twice the number of through holes, and the 2-nd Betti number equals the number of voids. For an isosurface (or in general a 2-manifold) only the first three Betti numbers can be non-zero.

3 Algorithm

In this section, we describe an algorithm that detects the tunnels and pockets using the critical point structure of the distance function.

3.1 Sampling of Level Set

In order to successfully apply the critical point structure of the discrete approximation of the distance function h_Σ by h_P, we require a suitable discrete approximation of the level set. Apparently primal contouring (Marching Cubes [27]) and dual contouring [26] are good choices to extract a discrete approximation of the level set from the scalar volume. Although variants of these approaches have been researched extensively to produce a topologically consistent isosurface, the main disadvantage lies in the fact that the sampling of the extracted surface is oblivious to the feature. Note, we need a set of points P to approximate h_Σ by h_P, and we also need h_P to follow h_Σ closely so that we do not miss the topological features of Σ in this process of translating it to the discrete setting. Recently, we have developed an algorithm which ensures that the discretization of the level set has sufficiently dense sampling for it to be a subcomplex of the Delaunay triangulation of the set of samples. This guarantee that the sampling is feature-sensitive and therefore the discretization follows closely the distance function induced by the true level set. Due to space limitation, we omit the details of the algorithm here and refer the reader to [23].

3.2 Classification and Clustering of Critical Points

The critical points of h_P are detected by checking the intersection of the Voronoi and its dual Delaunay diagram of the point set P sampled from Σ. The critical points are primarily of three types depending on if the Voronoi/Delaunay object involved lies interior, exterior to Σ, or if the Voronoi object crosses Σ. The maxima can not lie on the surface and therefore they are only of two types - interior and exterior. The minima are sample points themselves and therefore they are always on Σ. The saddle points can be any of three types mentioned above.

We use C_2 to denote the set of index 2 saddles which is partitioned into three classes $C_{2,I}$, $C_{2,O}$ and $C_{2,S}$. The set of maxima is denoted as C_3 which is partitioned into two classes $C_{3,I}$ and $C_{3,O}$. Using the hierarchical nature, we build an *incidence* graph over $C_2 \cup C_3$ where an edge is formed between $c_{2,*}$ and $c_{3,*}$ if stable manifold of $c_{2,*}$ is on the boundary of the stable manifold of $c_{3,*}$. The edges are colored depending on if $c_{3,*} \in C_{3,I}$ (red) or $\in C_{3,O}$ (green). The graph is further augmented by the edges within C_2 (blue) if two index 2 saddles' stable manifold have non-empty intersection. For the sake of compactification, we also need to consider the point at infinity which acts as an infinite maximum (m_∞) and therefore is an element of $C_{3,O}$.

We are now equipped with a well-defined structure over the set $C_2 \cup C_3$ which leads to a natural way of clustering the elements in the graph following the hierarchical nature of the stable manifolds. We employ the following three rules to perform the clustering. The rules are applied only on the subsets $C_{2,O}$, $C_{2,S}$ and $C_{3,O} \setminus \{m_\infty\}$.

- **Rule 1**: Two index 2 saddles $c_i, c_j \in C_{2,O}$ are in the same cluster if there is a blue edge between them.
- **Rule 2**: Two maxima $m_i, m_j \in C_{3,O} \setminus \{m_\infty\}$ are in the same cluster if there is a common index 2 saddle c_k which is connected to both m_i and m_j via green edge.
- **Rule 3**: Two index 2 saddles $c_i, c_j \in C_{2,S}$ are clustered together if they each have a green edge to possibly two different maxima $m_i, m_j \in C_{3,O}$ where both m_i, m_j are in the same cluster by Rule 2.

3.3 Detection of Tunnels and Pockets

These three rules produce a clustering of the set $C_{2,O} \cup C_{2,S} \cup C_{3,O} \setminus \{m_\infty\}$. Every cluster is then examined more closely in order to bring out finer invariant features. The index 2 saddles falling in a single cluster can again be of three types as enumerated below.

- Type A: If the stable manifold of an index 2 saddle point is at the boundary of two finite maxima, both from the set $C_{3,O}$.
- Type B: If the stable manifold of an index 2 saddle point is incident upon m_∞ and a single finite maximum from the set $C_{3,O}$.
- Type C: If the stable manifold of an index 2 saddle point is at the boundary of no finite maximum.

The index 2 saddles of type B or type C whose stable manifolds share a common boundary are collected together to form sub-clusters. The combined stable manifold of each such sub-cluster gives a polygonal patch, called *mouth*.

The number of *mouth*s helps detect the following topological features.

- 0 **Mouth** indicates that the cluster belongs to a *void*.
- 1 **Mouth** indicates that the cluster belongs to a *pocket*.
- $k \geq 2$ **Mouths** indicate the cluster belongs to a *tunnel*. We call it a k-mouthed tunnel.

We use the algorithms described in [21] for computation of the stable manifolds of index 2 saddles. In order to have a computational description of the detected features, we also compute the stable manifolds of the maxima falling into every cluster using the algorithm described in [16]. This produces a tetrahedral decomposition of the features captured. Figure 3 illustrates this process.

3.4 Ranking and Selective Removal of Tunnels and Pockets

The tetrahedral solids describing the pockets and tunnels provide a nice handle to those features and using these handles, the features can be ranked. We primarily use the geometric attributes of the features in order to rank them. Such attributes include, but are not limited to, the combined volume of the tetrahedra and the area of the mouths. The pockets and tunnels are then sorted in order of their increasing geometrically measured importance.

Removal of insignificant features are also made easy because of the volumetric description of the features. As dictated by the applications, a cut-off is set below which the features are considered *noise*. We remove the *topological noise* by marking the outside tetrahedra as inside and updating the surface triangles.

4 Results

We show the results of our algorithm on two volumetric data. The top row in Figure 4 shows the electron density volume of Rieske Iron-Sulfur Protein (Protein Data Bank Id: 1RIE). The volume rendering using VolRover [13] is shown in the leftmost subfigure. The tool additionally supports the visualization and isosurface selection using CT. The other subfigures show the selected interface and the detected tunnels and pockets. Note, the mouth of the tunnel is drawn in red and the mouth of the pocket is drawn in purple. The rest of the tunnel surface is drawn in yellow while the pocket surface is drawn in green. The blue patches in the rightmost subfigure shows the filling of the smaller tunnels and pockets. The second row shows the results on the three dimensional scalar volume representing the electron density of the reconstructed image of the chaperonin GroEL from a set of two dimensional electron micrographs. The resolution is $8\mathring{A}$. Using VolRover, a suitable level set is chosen. Note the CT is very noisy and has many branches, because of which it is not possible to extract a single-component isosurface. Nevertheless only one component is vital and the rest of them are merely artifacts caused by noise. The main component along with the detected tunnel is shown next. The result is particularly useful in visualizing the symmetric structure of the chaperonin as depicted in the symmetric set of mouths. In addition to detecting the principal topological feature, the algorithm detects few small tunnels and pockets which are shown separately for visual clarity (rightmost subfigure) and these are removed subsequently as part of the topological noise removal process.

We must also mention that, the presented approach for curation can also be applied to the modeling of smaller subunits of macromolecular complex, like viruses. In such cases, the complex is first segmented from into its building blocks [33] and they are structurally analyzed either via image processing [34] or via geometry processing [2]. A comprehensive survey on available computational approaches for modeling biological entities from electron density maps can be found in [3].

5 Conclusion

In this paper, we have presented an algorithm which, given an isosurface extracted from a scalar volume, captures the topological and geometric characteristics of the

isosurface and allows for the selective removal of unwanted features. The strength of the algorithm lies in its ability to connect the topology of the level set with the critical point structure of the distance function induced by the level set.

Acknowledgment: This research was supported in part by NSF grants IIS-0325550, CNS-0540033 and NIH contracts P20-RR020647, R01-9M074258, R01-GM07308, R01-EB004873.

References

1. Bajaj, C., Bernardini, F., and Xu, G. Automatic reconstruction of surfaces and scalar fields from 3D scans. In *ACM SIGGRAPH* (1995), pp. 109–118.
2. Bajaj, C., and Goswami, S. Automatic fold and structural motif elucidation from 3d EM maps of macromolecules. In *ICVGIP 2006* (2006), pp. 264–275.
3. Bajaj, C., and Yu, Z. Geometric and signal processing of reconstructed 3d maps of molecular complexes: Handbook of computational molecular biology. Computer and Information Sciences Series. Chapman and Hall / CRC press, December 21, 2005.
4. Bajaj, C. L., Pascucci, V., and Schikore, D. R. The contour spectrum. In *Proceeding Visualization '97* (Phoenix, AZ, October 1997), R. Yagel and H. Hagen, Eds., IEEE Computer Society & ACM SIGGRAPH, pp. 167–173.
5. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., Shindyalov, I., and Bourne, P. The Protein Data Bank. *Nucleic Acids Research* (2000), 235–242.
6. Boyell, R., and Ruston, H. Hybrid techniques for real-time radar simulation. In *Fall Joint Computer Conference '63* (Las Vegas, NV, 1963), pp. 445–458.
7. Carlsson, G., Zomorodian, A., Collins, A., and Guibas, L. J. Persistence barcodes for shapes. *International Journal of Shape Modeling 11*, 2 (2005), 149–187.
8. Carr, H., Snoeyink, J., and Axen, U. Computing contour trees in all dimensions. *Computational Geometry Theory and Applications 24* (2001), 75–94.
9. Carr, H., Snoeyink, J., and van de Panne, M. Simplifying flexible isosurfaces using local geometric measures. *IEEE Visualization 2004* (2004).
10. Chaine, R. A geometric convection approach of 3D reconstruction. In *Proc. Eurographics Sympos. on Geometry Processing* (2003), pp. 218–229.
11. Chazal, F., and Lieutier, A. Stability and homotopy of a subset of the medial axis. In *Proc. 9th ACM Sympos. Solid Modeling and Applications* (2004), pp. 243–248.
12. Collins, A., Zomorodian, A., Carlsson, G., and Guibas, L. A barcode shape descriptor for curve point cloud data. *Computers and Graphics 28* (2004), 881–894.
13. CVC, UT Austin. Volrover. *http://cvcweb.ices.utexas.edu/software/guides.php*.
14. de Berg, M., and van Kreveld, M. Trekking in the alps without freezing or getting tired. *Algorithmica 18*, 3 (1997), 306–323.
15. de Silva, V., and Carlsson, G. Topological estimation using witness complexes. In *Point-Based Graphics '04* (2004), M. Alexa and S. Rusinkiewicz, Eds., pp. 193–199.

16. Dey, T. K., Giesen, J., and Goswami, S. Shape segmentation and matching with flow discretization. In *Proc. Workshop Algorithms Data Strucutres (WADS 03)* (Berlin, Germany, 2003), F. Dehne, J.-R. Sack, and M. Smid, Eds., LNCS 2748, pp. 25–36.
17. Edelsbrunner, H. Surface reconstruction by wrapping finite point sets in space. In *Ricky Pollack and Eli Goodman Festschrift*, B. Aronov, S. Basu, J. Pach, and M. Sharir, Eds. Springer-Verlag, 2002, pp. 379–404.
18. Edelsbrunner, H., Facello, M. A., and Liang, J. On the definition and the construction of pockets in macromolecules. *Discrete Apl. Math. 88* (1998), 83–102.
19. Edelsbrunner, H., Letscher, D., and Zomorodian, A. Topological persistence and simplification. *Discrete Comput. Geom. 28* (2002), 511–533.
20. El-Sana, J., and Varshney, A. Controlled simplification of genus for polygonal models. In *Proceedings of the IEEE Visualization '97* (1997), pp. 403–412.
21. Giesen, J., and John, M. The flow complex: a data structure for geometric modeling. In *Proc. 14th ACM-SIAM Sympos. Discrete Algorithms* (2003), pp. 285–294.
22. Goswami, S., Dey, T. K., and Bajaj, C. L. Identifying flat and tubular regions of a shape by unstable manifolds. In *Proc. 11th ACM Sympos. Solid and Phys. Modeling* (2006), pp. 27–37.
23. Goswami, S., Gillette, A., and Bajaj, C. Efficient Delaunay mesh generation from sampled scalar function. In *Proc. 16th Int. Meshing Roundtable* (2007), p. to appear.
24. Guillemin, V., and Pollack, A. *Differential Topology.* Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1974.
25. Guskov, I., and Wood, Z. Topological noise removal. In *Graphics Interface* (2001), pp. 19–26.
26. Ju, T., Losasso, F., Schaefer, S., and Warren, J. Dual contouring of hermite data. In *Proceedings of SIGGRAPH* (2002), pp. 339–346.
27. Lorensen, W., and Cline, H. Marching cubes: A high resolution 3d surface construction algorithm. In *ACM SIGGRAPH '87* (1987), pp. 163–169.
28. Nooruddin, F., and Turk, G. Simplification and repair of polygonal models using volumetric techniques. Tech. Rep. Research Report 99-37, Georgia Tech., 1999.
29. Pascucci, V., and Cole-McLaughlin, K. Parallel computation of the topology of level sets. *Algorithmica 38* (2003), 249–268.
30. Siersma, D. Voronoi diagrams and morse theory of the distance function. In *Geometry in Present Day Science*, O. E. Barndorff and E. B. V. Jensen, Eds. 1999, pp. 187–208.
31. van Kreveld, M., van Oostrum, R., Bajaj, C., Pascucci, V., and Schikore, D. Contour trees and small seed sets for isosurface traversal. In *Proc. 13th ACM Symposium on Computational Geometry* (1997), pp. 212–220.
32. Wood, Z., Hoppe, H., Desbrun, M., and Schroder, P. Removing excess topology from isosurfaces. *ACM Transactions on Graphics 23*, 2 (April 2004), 190–208.
33. Yu, Z., and Bajaj, C. Automatic ultrastructure segmentation of reconstructed cryoem maps of icosahedral viruses. *IEEE Transactions on Image Processing: Special Issue on Molecular and Cellular Bioimaging 14*, 9 (2005), 1324–37.

34. Yu, Z., and Bajaj, C. Computational approaches for automatic structural analysis of large bio-molecular complexes. *IEEE/ACM Transactions on Computational Biology and Bioinformatics accepted for publication* (2007).
35. Zhang, X., Bajaj, C., Kwon, B., Dolinsky, T., Nielsen, J., and Baker, N. Application of new multiresolution methods for the comparison of biomolecular electrostatics properties in the absence of structural similarity. *Multiscale Modeling and Simulation 5*, 4 (2006), 1196–1213.
36. Zomorodian, A. *Computing and Comprehending Topology: Persistence and Hierarchical Morse Complexes*. PhD thesis, University of Illinois at Urbana-Champaign, 2001.
37. Zomorodian, A., and Carlsson, G. Computing persistent homology. *Discrete Comput. Geom. 33*, 2 (2005), 249–274.

Appendix: Figures and Captions

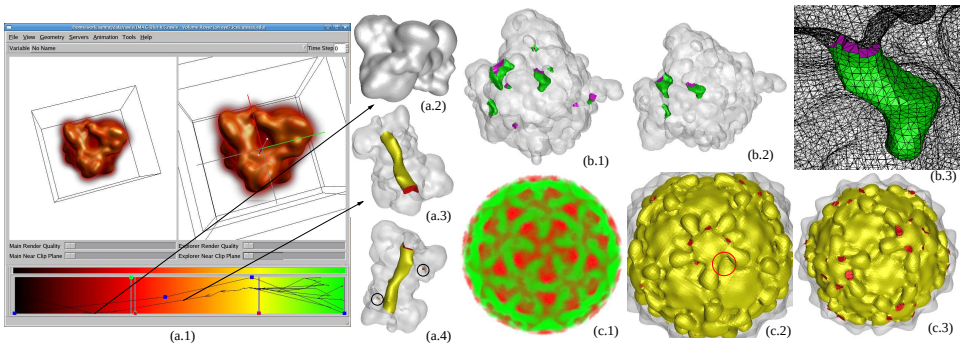

Fig. 1. Three examples of molecular surface selection are shown. (a.1-4) shows the selection of Gramicidin A that preserves the ion channel. (b.1-3) shows the selection of molecular surface for mouse Acetylcholinesterase (mAchE) where the pocket near the active site for binding Acetylcholine is preserved. (c.1-3) shows the selection of isosurface of the viral capsid of Nodavirus that brings out the inherent symmetry. In all three examples, we show how our algorithm for detecting and evaluating topological features aids the proper isosurface selection.

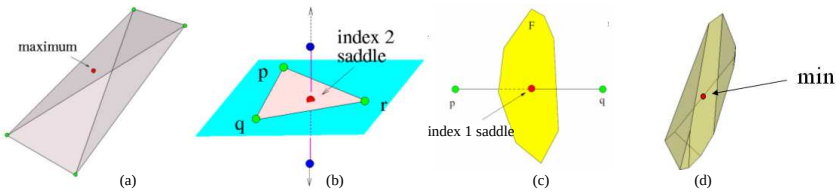

Fig. 2. The relative position of Voronoi and their dual Delaunay objects that results in the generation of critical points.

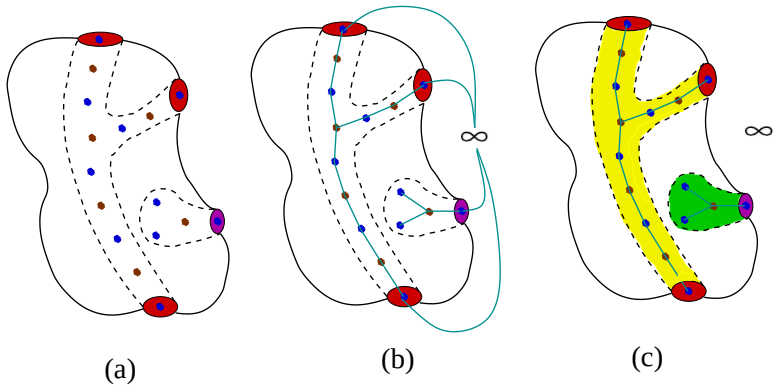

Fig. 3. An illustration of our tunnel and pocket detection algorithm. An imaginary molecular surface is shown with a 3-mouth tunnel and a single pocket. (a) Critical points of h_P are detected. Blue points are index 2 saddles and brown points are maxima. (b) A point at infinity is added and critical points are clustered based on adjacency of stable manifolds. (c) Based on the saddle points incident on infinity, we detect and classify the tunnel (yellow with red mouths) and pocket (green with purple mouth).

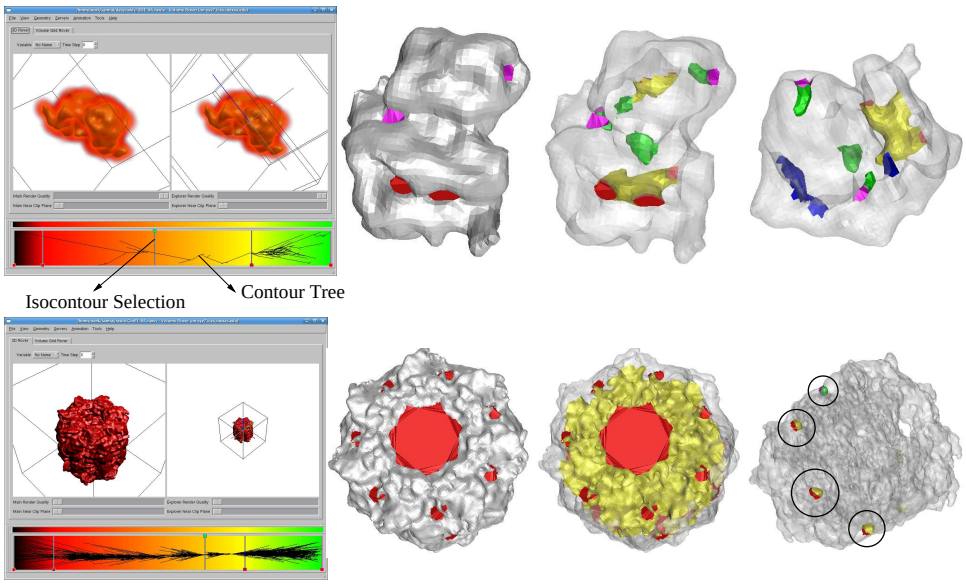

Fig. 4. Results: Top row shows the interface selection for Rieske Iron-sulfur Protein molecule (PDB ID: 1RIE) from a blurred density map. Bottom row shows the isosurface selection for the chaperonin GroEL from cryo-EM density map.