Topology-Controlled Volume Rendering

Gunther H. Weber, Member, IEEE Computer Society, Scott E. Dillard, Hamish Carr, Member, IEEE, Valerio Pascucci, Member, IEEE, and Bernd Hamann, Member, IEEE

Abstract—Topology provides a foundation for the development of mathematically sound tools for processing and exploration of scalar fields. Existing topology-based methods can be used to identify interesting features in volumetric data sets, to find seed sets for accelerated isosurface extraction, or to treat individual connected components as distinct entities for isosurfacing or interval volume rendering. We describe a framework for direct volume rendering based on segmenting a volume into regions of equivalent contour topology and applying separate transfer functions to each region. Each region corresponds to a branch of a hierarchical contour tree decomposition, and a separate transfer function can be defined for it. The novel contributions of our work are 1) a volume rendering framework and interface where a unique transfer function can be assigned to each subvolume corresponding to a branch of the contour tree, 2) a runtime method for adjusting data values to reflect contour tree simplifications, 3) an efficient way of mapping a spatial location into the contour tree to determine the applicable transfer function, and 4) an algorithm for hardware-accelerated direct volume rendering that visualizes the contour tree-based segmentation at interactive frame rates using graphics processing units (GPUs) that support loops and conditional branches in fragment programs.

Index Terms—Direct volume rendering, transfer function design, topology, contour tree, simplification.

1 INTRODUCTION

Volume rendering is a standard technique used in scientific visualization. It is based on defining “optical” properties for points in the three-dimensional (3D) domain of a scalar field and computing resulting pixel intensities in an image plane [1]. Optical properties at a given point are normally based wholly or partially on the function value (and sometimes gradient magnitude). This approach assumes that scalar values, such as density values in a computed tomography (CT) scan, map directly to physical properties such as tissue type. Even when volume rendering is applied to nonmedical data this model is often used [2], [3]. While this approach provides a structural overview of an entire data set, it is unable to distinguish between distinct features that share the same scalar value. If, for example, a region of interest is enclosed by an “uninteresting” region and both regions overlap in value range, occlusion generally prevents effective visualization of the interesting region.

Improved methods for defining transfer functions utilize additional derived quantities such as gradient magnitude [4], but they still apply the same transfer function uniformly throughout the domain. As a result, these methods, while effective at enhancing the perception of distinct material/tissue type interfaces, still do not allow a viewer to differentiate fully between separate features. Other methods use segmentation information to apply different transfer functions to classified regions [5], [6] or even to apply different rendering modalities [7], [8]. These techniques require segmentation information for each volume and rely either on manual segmentation or domain-specific segmentation algorithms.

The underlying weakness of all of these methods is that the determination of opacity properties either fails to distinguish between different features at similar isovalues, or is based on expensive and/or domain-specific segmentation methods. It is therefore desirable to have a general model of spatially local transfer function definition by exploiting topological definitions of what constitutes a feature. By incorporating topological information into direct volume rendering in a systematic fashion, we generalize existing topological inputs to transfer function design and provide methods that are practical for direct user manipulation or, eventually, automatic design of locally defined transfer functions.

Isosurface topology provides insight into the fundamental structure of a data set, independently of the application domain. Topology considers connected components in a relatively intuitive fashion that can be used for a wide variety of applications. The contour tree [9] provides expert users with additional information about their data and facilitates data exploration in the absence of strong domain-specific techniques. Topological analysis of scalar data sets can be used to identify interesting behavior of a scalar field and to aid in the data exploration process. Section 2 provides an overview of related work.

While specialized segmentations are applicable to particular problem domains, the power of our topology-based segmentation lies in its generality. Our approach is related to the work of Takeshima et al. [10], who used a topological abstraction called volume skeleton tree to define topological attributes as additional input for multidimensional transfer
functions. Our method extends their ideas and is more general. Instead of setting transfer functions according to fixed topological indices, such as depth of topological nesting, we allow a user to assign independent transfer functions to topologically distinct features which may or may not share the same indices. In contrast to the work of Takeshima et al., we define the opacity at a sample point based on its topological characterization rather than interpolating between the opacities at corners of the associated grid cell. In essence, our approach performs postclassification in volume rendering while Takeshima et al.’s approach uses preclassification. While the use of postclassification for medical data is controversial, it significantly reduces artifacts in rendered images [11] and is particularly suitable for simulated data. Furthermore, instead of using linear interpolation applied to a tetrahedron consider trilinear interpolation for a hexahedral cell, as hexahedral meshes are most commonly used for direct volume rendering of data given as samples on a regular, rectilinear mesh.

We have implemented topological data simplification for trivariate scalar data sets. Topological data simplification techniques simplify a data set by making its topology correspond to hierarchical simplifications of the contour tree, thus representing various levels of detail and showing significant features at a glance while still allowing fine details to be revealed on demand. Existing simplification methods result in geometric and visual artifacts such as fine thread-like structures and flat regions showing up as large objects abruptly becoming visible. We demonstrate how to remove topological features in a way that reduces these problems (Section 4). We also describe data structures (Section 3) and rendering algorithms for topological volume rendering (Section 5) that can be implemented efficiently using current graphics hardware. Our framework supports interactive volumetric visualization of noisy or topologically complex data.

2 RELATED WORK
2.1 Contour Trees
In addition to its importance for isosurface extraction [12], topology has become valuable for more general exploration of scalar fields. Topology, up to now, has been used mainly in the context of isosurface extraction, where the topological properties of interest are usually the number of connected components and the genus of an isosurface, i.e., the number of independent tunnels or “holes” in an isosurface. Morse theory [13] shows that topological changes in scalar fields defined on manifolds occur at distinct isolated points called critical points. A Reeb graph [14] expresses the evolution of individual contours as a graph that is defined by these critical points and their relationships. For simply connected domains, the Reeb graph is always a tree structure, called a contour tree [9], which is more easily computed than the general Reeb graph.

Before discussing the contour tree in detail, we start with some basic definitions. The term isosurface has been used in the literature to mean the inverse image of an isovalue, an individual connected component of such an inverse image, or a triangular approximation of either. In mathematics, the term level set is usually used, instead of isosurface, but this term has other connotations in graphics and visualization. To avoid confusion, we use the terms isosurface and contour as follows: Given a field \( f : \mathbb{R}^3 \to \mathbb{R} \), the isosurface of \( f \) for an isovalue \( h \) is the inverse image \( f^{-1}(h) \) of the isovalue. Since this isosurface may consist of multiple connected components, we use contour to refer to an individual connected component of an isosurface.

The contour tree is a structure that captures the topological evolution of an isosurface as the isovalue varies. Its nodes are critical points of the data set where the number of contours changes. Nodes of degree one (leaves of the tree) are minima and maxima where contours are created or destroyed. Interior nodes of degree three or higher are saddles where two or more contours merge or a single contour separates into multiple disconnected contours. Arcs of the contour tree represent contours between critical points, i.e., contours which do not change topology (with the exception of genus changes) as the isovalue varies between critical values. Algorithms for computing the contour tree [15], [16], [17], [18] were first defined for tetrahedral meshes using linear interpolation and later for hexahedral meshes using trilinear interpolation [19].

The contour tree may be augmented by additional nodes representing changes in topological genus of contours [19], nodes representing noncritical vertices of the mesh, or nodes representing any arbitrary points in the volume. We use the term fully augmented contour tree to refer to the contour tree augmented by all vertices of a mesh. Each grid vertex corresponds to a node in the fully augmented contour tree and maps to an arc (when a regular point) or a node (being a critical point) in the contour tree.

Since each contour in an isosurface corresponds one-to-one with a point on an arc in the contour tree, it is possible to use the contour tree as an index for a volume data set and identify all contours for a given isovalue. This relationship has also been used to find seed cells [17], [20] for the continuation method [21] of level set extraction.

2.2 Utilizing Topology in Scientific Visualization
Even if the topology is known, a user interface for effective data exploration must still be developed. Bajaj et al. [22] introduced an interface called the contour spectrum, where properties such as isosurface area, enclosed volume, and the contour tree were plotted alongside isosurfaces to provide users with additional cues to interesting isovalue changes. Weber et al. [23], [24] devised tools for exploring scalar fields based on detecting critical points and critical regions. Cox et al. [25] defined a digital Morse theory and also used critical points and regions to explore medical data. Fujishiro et al. [26], [27] used the contour tree to detect significant isovalue changes automatically for transfer function design.

Carr and Snoeyink [20] extended the idea of seeded isosurface extraction based on the contour tree [17] by associating seed cells with individual contours, using the contour tree as a visual index to the contours. This association underpins the flexible isosurface interface in which individual contours are treated as distinct entities. Since this concept supports independent manipulation of single contours, it imports a notion of spatial locality to the task of extracting surfaces from a scalar field. Individual contours can be deleted (allowing a user to view otherwise
as “peeling” off layers in a volumetric data set. Dimensional transfer functions, to support operations such as isosurface inclusion level in multi-dimensional volumes to display volumetric regions of uniform topology. Recently, Takeshima et al. [10] used topological “peeling” away layers of a volume to examine the internal structure. Noise in a data set creates a large number of irrelevant critical points that can distract from truly relevant topological features. Topology simplification suppresses insignificant features by removing, or cancelling, pairs of critical points that are viewed unimportant according to a specified measure. Besides the contour tree, two topological structures are used widely for scalar topology simplification: the volume skeleton tree that corresponds to a contour tree augmented with nodes corresponding to genus changes [29], [30] and the Morse-Smale complex [31], [32], [33].

Our work is based on topology simplification methods introduced by Takahashi et al. [28], Carr et al. [34], and Pascucci et al. [35]. Carr et al. [34] simplified the contour tree with two basic operations: leaf pruning and node collapses. Leaf pruning removes a leaf and the arc incident to the leaf from the contour tree. Since contours are extracted from seeds stored in a contour tree, deleting an arc from the contour tree has the side effect of discarding the corresponding contours from further consideration. Carr et al. further show that when visualization methods other than isosurface extraction are used, data can be modified to match the topology of the simplified contour tree by “flattening” the corresponding region, i.e., all points corresponding to the pruned arc get assigned the value of the saddle. Node collapses remove degree-two vertices and do not affect the contours that can be extracted or values in the data set. Pruning and collapsing are performed in an order that minimizes the error based on a local geometric measure, such as hypervolume (an integral of the scalar field over the enclosed volume), volume or persistence, with node collapses having priority over leaf pruning.

Pascucci et al. [35] described branch decomposition, an efficient way for storing a hierarchy of contour tree simplifications. A branch decomposition of a tree is a hierarchical decomposition if 1) exactly one branch connects two leaves (called the root branch) and 2) every other branch connects one leaf to an interior node of another branch. Fig. 1b shows a contour tree, and Fig. 1d shows its branch decomposition. Since a branch (with the exception of the root branch) connects a leaf to an interior node of another branch, it is defined by a pair of critical points: a saddle and an extremum that are connected by a monotone path. Each saddle-extremum pair corresponds to a topological simplification, or cancellation, of critical points. Discarding a branch from the branch decomposition is equivalent to performing a cancellation, which is equivalent to a vertex prune operation in the framework of Carr et al. [34].

2.3 Topology Simplification

Noise in a data set creates a large number of irrelevant critical points that can distract from truly relevant topological features. Our work is based on topology simplification methods introduced by Takahashi et al. [28], Carr et al. [34], and Pascucci et al. [35]. Carr et al. [34] simplified the contour tree with two basic operations: leaf pruning and node collapses. Leaf pruning removes a leaf and the arc incident to the leaf from the contour tree. Since contours are extracted from seeds stored in a contour tree, deleting an arc from the contour tree has the side effect of discarding the corresponding contours from further consideration. Carr et al. further show that when visualization methods other than isosurface extraction are used, data can be modified to match the topology of the simplified contour tree by “flattening” the corresponding region, i.e., all points corresponding to the pruned arc get assigned the value of the saddle. Node collapses remove degree-two vertices and do not affect the contours that can be extracted or values in the data set. Pruning and collapsing are performed in an order that minimizes the error based on a local geometric measure, such as hypervolume (an integral of the scalar field over the enclosed volume), volume or persistence, with node collapses having priority over leaf pruning.

Pascucci et al. [35] described branch decomposition, an efficient way for storing a hierarchy of contour tree simplifications. A branch decomposition of a tree is a hierarchical decomposition if 1) exactly one branch connects two leaves (called the root branch) and 2) every other branch connects one leaf to an interior node of another branch. Fig. 1b shows a contour tree, and Fig. 1d shows its branch decomposition. Since a branch (with the exception of the root branch) connects a leaf to an interior node of another branch, it is defined by a pair of critical points: a saddle and an extremum that are connected by a monotone path. Each saddle-extremum pair corresponds to a topological simplification, or cancellation, of critical points. Discarding a branch from the branch decomposition is equivalent to performing a cancellation, which is equivalent to a vertex prune operation in the framework of Carr et al. [34].

2.4 Classification in Volume Rendering

Original approaches for volume rendering consider only the scalar value to classify samples [1], [3]. Even though Levoy’s method [1] also considers gradient magnitude it is only used to simulate “surfaces of constant thickness” and does not affect classification. Kniss et al. [4] improved visualizations by adding gradient magnitude and additional higher-order derivatives as parameters for classification, while still applying the same transfer function uniformly throughout the domain.

If a segmentation of the data set is available, which is often true for medical data sets, it is possible to improve volume-rendered results by considering segmentation information during classification and rendering individual segmented regions differently. For example, it is possible to use different transfer functions for different tissue types, or to emphasize particular parts of the anatomy while hiding others [5], [6]. By combining multiple rendering modalities, such as maximum intensity projection and direct volume rendering with incorporation of different transfer functions [7], [8], it is possible to highlight volume portions of interest while still providing a context to the user. However, these techniques require segmentation information which must be obtained by manual segmentation or by domain-specific segmentation algorithms.
More recently, Takeshima et al. [10] used information from the contour tree to design multidimensional transfer functions which use derived topological quantities as input. In addition to genus, the authors designed transfer functions which express the “inclusion level” of a contour, defined as the number of equal-valued contours which surround and occlude that contour. Lower inclusion levels, which represent the outermost features, are given lower opacities so that higher inclusion levels, which represent the innermost features can be seen clearly. Furthermore, Takahashi et al. [36] used the volume skeleton tree to find optimal viewpoints that maximize the number of visible features for volume rendering.

Recently, statistical learning-based methods have been introduced for interactive volume data segmentation [37]. The idea underlying such methods is to have a user specify interactively what regions in a data set constitute a “feature.” By pointing out such regions, it is then possible to characterize them by scalar field behavior in a local neighborhood, and to use the resulting characterization for segmentation. However, learning-based methods require manual training to define features.

Kniss et al. [38] improved on the classical classification by observing that a unique classification of samples is not always possible. In particular, in the presence of boundary noise, it improves visualization results by considering an uncertainty interval if classification and use statistical methods to generate “fuzzy” classifications.

In summary, a significant body of work exists concerned with the use of contour trees for interactive isosurface exploration, on their use in automatic transfer function design, and in designing globally uniform transfer functions based on parameters which may include some topological information. However, interactive exploration of scalar fields using topologically defined transfer functions in their most general sense has not previously been considered, and our paper discusses a solution to this problem.

3 CONTOUR-TREE-DEFINED SEGMENTATION

Since individual contours map to points in the contour tree, each arc of the tree represents the union of all contours which map to it. This union can be thought of as the volume being swept out by the contour as its iso-value is varied, starting at the critical value which creates the contour, and ending at the value which destroys it. This sweep defines a partition of the space into topologically distinct regions, which we refer to as topological zones, following Cox et al. [25]. For example, in Fig. 1a, the sample terrain has been divided along each contour that passes through a critical point, and the topological zones are shown in different colors.

We also follow Takeshima et al. [10] by using topological zones for segmentation in volume rendering. However, their work has several limitations: their topological indices treat all zones with the same nesting depth as equivalent, ruling out differentiation between two topologically similar but distinct objects. Moreover, their methods are based on the topology of tetrahedral meshes, which can introduce geometric and topological artifacts in the output images [39]. Furthermore, they computed opacities for the volume rendering integral by interpolating over tetrahedra, instead of looking up opacities directly in the relevant transfer function, which can lead to further artifacts in cells spanning multiple topological zones.

We extend this work by allowing assignment of independent transfer functions to individual topological zones; by applying the trilinear interpolant over hexahedral cells; by showing how to find the correct opacity in cells spanning multiple topological zones; and by improving the treatment of topological simplification in the data.

We compute the contour tree of the underlying trilinear interpolant [12] with a variation of the method from Pascucci and Cole-McLaughlin [19], which applied the underlying graph-theoretic algorithm [18] to the union of cell-wise join and split trees instead of the edges of a tetrahedral mesh. In fact, any graph which accurately represents the topology of the trilinear interpolant may be substituted instead. To simplify processing, we construct a widget in each cube in which the mesh edges are augmented by edges between body saddles, face saddles, and vertices [40]. Nothing, however, depends on this choice, and the method described in [19] can be substituted instead.

Once we have computed the contour tree, we simplify it to remove symbolic perturbation [41], then assign each topological zone a transfer function. Consequently, every topological detail, including noise, is assigned its own transfer function. For medium to large data sets, especially when physically sampled, the topological complexity makes it necessary to represent the contour tree hierarchically. We store the hierarchy using the branch decomposition described by Pascucci et al. [35].

In addition to providing a hierarchical representation of the contour tree, branch decomposition also avoids an “overssegmentation” of the volume. While having a topological zone per contour family with equivalent topology is useful, it does not take into account that a segmentation is only necessary if regions overlap in value range. Branch decomposition naturally concatenates a sequence of arcs into a single unit, see Fig. 1d, unifying their topological zones into a single zone, see Fig. 1c. Furthermore, branch decomposition permits user interactions on a coarse, simplified tree and propagating transfer functions down to the full resolution tree.

We construct the branch decomposition as follows: At each step, a saddle-extremum pair is pruned from the contour tree, and the remaining node is collapsed, leaving an arc. This arc is the parent to which the saddle-extremum pair is linked as a child. The order in which the pairs are pruned is determined by a priority measure. Pascucci et al. [35] used persistence as a priority measure, which is the absolute difference in function values between the saddle and extremum. Geometric measures, such as volume and hypervolume, are described by Carr et al. [34]. Pairs with lower priority are pruned first, and they become the children of pairs with higher priority. The end result is a rooted tree of saddle-extremum pairs, or branches. The most important branches, as indicated by the priority measure, are at the top of this tree.

Currently, we do not consider saddles corresponding to genus changes. Our zones of “topologically equivalent” contours are only topologically equivalent up to the genus. However, the same argument for using a branch decomposition to avoid an overssegmentation also holds here. Saddles which correspond to a genus change of a contour do not affect the segmentation—contours of different genus can be distinguished based solely on function value alone.
This is a design decision, however, and nodes for genus change could be represented if the additional information were desired. As genus changes do not contribute to the complexity of the contour tree-based segmentation, we do not consider them in our simplification scheme.

4 Improved Topological Simplification

One straightforward way to change isosurface topology is to “chop off” peaks, and correspondingly “fill in” valleys [34]. (Figs. 2a and 2b show an example, where the peak on the right-hand side (marked C) is chopped off.) This goal can be achieved by setting the transfer function for the canceled branch to a solid color, the color of the parent branch’s transfer function at the value of the saddle which connects the child to the parent. While this approach displays an image with topology consistent with the pruned contour tree, the topological changes in the function may not be ideal from a user’s perspective. Sweeping an isosurface down the simplified contour tree will show the pruned components “popping” abruptly into existence, as can be seen in the movie accompanying this paper (which can be found at http://www.computer.org/tvcg/archives.htm). In volume-rendered images, a region of uniform opacity properties can also become visible. For a smoother evolution of simplified isosurfaces, or for a closer approximation of the underlying scalar field, other methods must be used.

We follow the strategy of Pascucci et al. [35], which proves that one can always modify a scalar field to correspond to a simplified contour tree. The proof relies on reversing the gradient flow along a path from the saddle to the extremum of the branch to be pruned. In fact, the authors proved a stronger result, which states that the region in which the gradient flow is reversed can be made arbitrarily small. This condition is achieved by subdividing the mesh, resulting in arbitrarily thin thread-like structures connecting the former extremum to the remainder of the function. This approach, too, generates undesirable visual artifacts, and we instead reverse the gradient flow in a tube—a region inflated from the thread to a size defined by the following heuristic: sweep a sphere out from the saddle (S in Fig. 2). The base radius of the tube is the distance at which the sphere first intersects a topological zone other than the parent’s or child’s. This radius shrinks as the tube extends from the saddle toward the extremum.

Consider a branch that ends at a maximum, such as the branch SC in Fig. 2. The region of gradient reversal is a tube, the child tube, that starts at the saddle S and extends “upward” toward the maximum C. This child tube crosses each contour of SC’s topological zone once, and at each crossing we define a new function value. The original values of the contours ascend from the saddle to the maximum. To reverse the gradient, we define new values that descend, as illustrated in Fig. 2c. Since the starting point of this tube, S, is lower than all values in the topological zone of C, the path cannot descend without creating a new minimum. To avoid creating this new minimum, we increase the value at S, by way of a parent tube which extends from the saddle into its parent’s topological zone, PS in Fig. 2. The gradient along this parent tube is not reversed. Instead, the values are increased in order to elevate the saddle’s function value. If the value at P is greater than the value at C (which is usually true for the hierarchical branch decomposition [35]) then it is possible to construct a monotonically decreasing path from P, through S, ending at C, thus effecting the simplification.

The tubes are created by a “best effort” process, which starts at the saddle and follows the steepest path in both directions. It stops if a maximum or minimum is reached, or if the path is obstructed by another topological zone. If the end of the parent tube does not have a value greater than the value at the end of the child tube, then a monotonically descending path cannot be constructed along the tube axis. In such a case, where the value at the highest point on the parent tube reaches a value v, we uniformly scale all values in SC’s topological zone so that they are less than v.

If the child SC has children, then the child tube may also become obstructed. It becomes necessary to flatten all points in C’s topological zone with value greater than at the end of the child tube. The tip of the child tube thus becomes the new maximum. In extreme cases where no significant tubes can be constructed, the method degenerates into the “chopping-off” method mentioned at the beginning of this section.

5 Volume Rendering

Volume rendering is based on integration of optical properties along a ray passing through the image plane. Samples are taken at regular intervals along this ray, and color and opacity are determined using a transfer function. Results are then composited. Various methods have been developed to accelerate this process, but all share the same conceptual model.

Instead of a global transfer function, we use a (possibly) distinct transfer function for each topological zone. For any given point p in the domain, we must therefore determine the topological zone to which p belongs, and then apply the appropriate transfer function. This task is complicated by the fact that the boundaries between topological zones are formed by isosurfaces at critical values, which rarely
coincide neatly with the mesh. Fig. 1a shows that the triangle labeled $A$ intersects three topological zones, the middle of which contains no vertex of $A$. In general, cells can intersect arbitrarily many topological zones.

Our solution to this problem is based on the fact that monotone paths in any scalar field always map to monotone paths in the contour tree and vice versa. Given a monotone path $P$ in the scalar field that passes through $p$ and terminates at vertices corresponding to nodes in the contour tree, we trace the corresponding monotone path $Q$ in the contour tree. Due to monotonicity, there will only be one point $q$ on this path with the same isovalue as $p$; thus, we apply the transfer function for the branch on which $q$ lies.

To determine a suitable monotone path, we observe that, for any point $p$, there always exists a monotone path $P$ through $p$ that starts at a local maximum and ends at a local minimum. Moreover, linear interpolation on tetrahedra and trilinear interpolation on hexahedra both guarantee that monotone paths in any scalar field always map to monotone paths in the contour tree, as shown in Fig. 3a. This fact simplifies our task, as we do not need to find a monotone path explicitly. Instead, we use the path from highest to lowest-valued vertices of the tetrahedron.

Trilinear interpolation on hexahedra does not have this property, since saddles may occur inside or on faces of hexahedral cells, causing monotone paths in the cell to map to more than one possible monotone path in the contour tree, as shown in Fig. 3b. We instead base our construction on the linearity of the interpolant along lines parallel to the coordinate system, and construct a monotone path from axis-parallel segments. Fig. 4 shows an example for the trilinear case.

A monotone path through a trilinear cell is constructed according to Algorithm 1. At any point $X$, we choose an axis-parallel line and follow it to the boundaries of the cell. Since $f$ is linear for every axis-parallel line, it must be either constant or monotone. If it is monotone, we repeat the process on the boundary of the cell until we reach a vertex, extending a monotone path as we proceed. For constant lines, the endpoints of the line belong to the same contour and the same topological zone as $X$, and we substitute either endpoint of the line for $X$ and continue.

Input $t : [0,1]^3 \rightarrow \mathbb{R}$ trilinear interpolant inside cell;
Sample position $p \in [0,1]^3$ inside trilinear cell
Output: $p_{\text{lo}}$ and $p_{\text{hi}}$ positions of mesh vertices (end points of the monotone path through $p$)

Create original monotone path to boundary faces:
$p_{\text{lo}} = \text{roundDown}(x,p)$
$p_{\text{hi}} = \text{roundUp}(x,p)$
if $t(p_{\text{lo}}) > t(p_{\text{hi}})$ then swap($p_{\text{lo}},p_{\text{hi}}$)

Complete descending path to mesh vertex:
$\text{temp} = \text{roundUp}(y,p_{\text{lo}})$
$p_{\text{lo}} = \text{roundDown}(y,p_{\text{lo}})$
if $t(p_{\text{lo}}) > t(\text{temp})$ then swap($p_{\text{lo}},\text{temp}$)
$\text{temp} = \text{roundUp}(z,p_{\text{lo}})$
$p_{\text{lo}} = \text{roundDown}(z,p_{\text{lo}})$
if $t(p_{\text{lo}}) > t(\text{temp})$ then swap($p_{\text{lo}},\text{temp}$)

Complete ascending path to mesh vertex:
$\text{temp} = \text{roundUp}(y,p_{\text{hi}})$
$p_{\text{hi}} = \text{roundDown}(y,p_{\text{hi}})$
if $t(p_{\text{hi}}) < t(\text{temp})$ then swap($p_{\text{hi}},\text{temp}$)
$\text{temp} = \text{roundUp}(z,p_{\text{hi}})$
$p_{\text{hi}} = \text{roundDown}(z,p_{\text{hi}})$
if $t(p_{\text{hi}}) < f(\text{temp})$ then swap($p_{\text{hi}},\text{temp}$)

Algorithm 1: Algorithm for finding vertices of a cell where the monotone path through a sample starts and ends. The function $\text{roundDown}(\text{axis},p)$ returns the position of $p$ moved to the “lower” boundary along the specified axis, i.e., it replaces the component corresponding to the axis with zero. Analogously, $\text{roundUp}(\text{axis},p)$ returns the position of $p$ moved to the “upper” boundary along the specified axis, i.e., it replaces the component corresponding to the axis with one.

Once the high and low vertices are determined, we look up the high and low arcs. We then walk from the high arc towards the low arc, stopping at the arc which contains the sampled value of $f$. Contour trees are “free trees” without a root. It is not immediately obvious which direction to walk if we are trying to move from one arc to another. Instead of trying to navigate the contour tree directly, we walk in the branch decomposition, which is a rooted tree. Suppose we are trying to find the path between branches $b$ and $c$. Since the tree is rooted, $b$ and $c$ must share a common ancestor $a$. 
The path from \( b \) to \( c \) goes from \( b \) to \( a \) and then back down to \( c \). This path is easily retrieved using the parent links of the branch decomposition. Further, we do not need to walk the entire path from \( b \) to \( a \) to \( c \) in order to find the branch we are looking for, the one which contains the sample point. Instead, we follow the paths from \( b \) to \( a \) and from \( c \) to \( a \) simultaneously. At each step, we walk up the parent link of the deeper of the two branches \( b \) and \( c \). To avoid overshooting when walking in the direction of increasing function value, we stop when the step we are about to take goes past a critical point (a saddle) with a higher value than the sampled point. Similarly, when walking in the direction of decreasing function value, if the next step goes past a saddle with a lower value than the sample, we stop. If neither of these conditions is ever met, the walk will terminate when both paths reach the common ancestor, i.e., the desired branch.

6 User Interface for Transfer Function Design

To define transfer functions and apply them to topological zones, we have extended the flexible isosurface interface [20], which allows for direct user selection and manipulation of individual contours using the standard metaphor of “select-and-operate.” Conceptually, we would like the user to select a topological zone directly in a rendered image by clicking the mouse in a zone of interest. To achieve this, we project a ray through the data originating from the corresponding pixel, and select the first “visible” topological zone. Visibility is determined by evaluating the volume rendering integral for a single ray passing through the volume and finding the first sample with an opacity above a certain threshold. For volume-rendered data, this approach is not always straightforward, especially when dealing with multiple layers of topological zones. Thus, we provide a slice view in which the user can select the topological zone’s intersection with that slice. In either case, once we have determined a sample (along the ray) or a pixel (in a slice), we use the fully augmented contour tree to determine the topological zone to which the sample (or pixel) belongs.

We also provide the ability to directly select arcs in the contour tree, as in the flexible isosurface interface [20]. For this purpose, we display the contour tree in one of two representations. A TOPORRERY view in Fig. 11c shows the branch decomposition contour tree in a three-dimensional radial layout [35] where height corresponds to function value. We also provide a more traditional two-dimensional view of the contour tree as shown in Fig. 11b. In this view, the contour tree is drawn based on the convention that the \( y \)-coordinate corresponds to function value. We have also found it useful to include a vertical transfer function editor (as opposed to the more traditional horizontal layout) next to the contour tree interface. By linking the \( y \)-coordinates of the contour tree interface and transfer function editor, contextual topological information is available to the user as they edit the transfer function.

7 Implementation

7.1 Hardware-Accelerated Volume Rendering

Volume rendering based on topological zone look-up can be implemented in graphics processing unit (GPU) fragment programs that support “for-loops,” such as those of NV40-class NVIDIA graphics cards. The algorithms outlined in Section 5, the monotone path search and branch look-up, are straightforward to implement in a high-level shader language like Cg. The image data is stored as a 3D texture to which trilinear interpolation is applied, and the vertex-to-branch map is stored as a 3D texture to which nearest-neighbor interpolation is applied. The output of this texture are references (texture coordinates) to an array of branches. These branch records are stored logically in a 1D array, but since the size of any one texture dimension is limited, it becomes necessary to wrap this array as a 2D texture. The fields for each branch record are: a reference to its parent, the value at its saddle, its depth in the rooted branch decomposition, and a reference to its transfer function. The transfer functions are stored in another logically 1D, physically 2D, array. Transfer function references are texture coordinates, to which a function value is added to offset the texture coordinate and select the proper color.

7.2 Visualization of Simplified Topology

Visualization of simplified topology is also implemented in graphics hardware. We render tubes between saddles and extrema dynamically, rather than modifying the scalar field beforehand. Each tube extends from a saddle to an extremum and is defined by a list of “tube points.” To achieve gradient reversal, we define the new function values to be increasing along the tube, where they were previously decreasing (see Section 4). In addition to the new function value, we also define the radius for each tube point, which varies the thickness of the tube.

The path for a tube is defined in a preprocessing step. Since the tubes are monotone paths, each tube point falls on a unique contour and has a unique function value. Consequently, during rendering, we can look up tube points from a table, indexed by function value, similar to the way transfer functions are used. For each function value \( v \), the associated position \( p_t(v) \), tube radius \( r_t(v) \), and “replacement function value” \( f_t(v) \), i.e., the value that replaces the original value at that location, are stored, see Fig. 5.

We use a blending function \( b \) around the axis of the tube,

\[
b = \frac{\sigma}{r_t(v)} \left( r_t(v) - \| p_s - p_t(v) \| \right).\]
where \( p_t \) is the location of the current sample (i.e., the fragment currently considered by the fragment program) and \( v \) is the function value at that point. The values \( p_t \) and \( r_t \) are read from the look-up table (texture) that defines the current tube. The blending function specifies to what degree the replacement function value is used instead of the original function value at the current location. If it is zero, the replacement value has no effect on the current location. If it is one, the replacement value is used instead of the current function value. Between a blending function value of zero and one, the final function value \( f_t(v_{in}) \) (i.e., the result of the fragment program) is obtained by linear interpolation between the original function value \( v_{in} \) at the current location and the corresponding replacement function value \( f_t(v_{in}) \), i.e.,

\[
v_{out} = \text{clamp}(1 - b)v_{in} + \text{clamp}(b)f_t(v_{in}),
\]

where

\[
\text{clamp}(x) = \begin{cases} 
0 & \text{if } x \leq 0, \\
 x & \text{if } 0 < x < 1, \\
1 & \text{if } 1 \leq x.
\end{cases}
\]

We restrict the computed value of \( b \) to the interval \([0, 1]\). The parameter \( \sigma \) defines the width of a tube by controlling how quickly \( b \) falls off as one moves away from the tube’s central axis. It is important that \( \sigma \) is larger than one; otherwise, \( b \) attains its maximum value only along the axis. A larger \( \sigma \) value, such as two, ensures that the interior volume of the tube with a \( b \) value of one is large enough to be visible during rendering. For each voxel, the closest tube in its topological zone is used to compute a modified function value. This closest tube is determined in a preprocessing step by checking the parent tube and all child tubes of that zone and identifying the tube which is closest according to the blending function \( b \). A reference to the appropriate tube is stored in a 3D texture and accessed during rendering.

In addition to these modifications, we may also need to scale the function values of an entire topological zone (as mentioned at the end of Section 4). Recall that if the parent tube cannot be extended to a point with greater (less) function value than the child maximum (minimum), then the function values of the child zone are scaled down (up) to permit a monotonic path from the parent extremum to the child. This scaling is a uniform 1D scaling transformation about the saddle value. We store the parameters of this transformation as a scale-and-bias pair, one for each branch, in the branch textures. The scaling is performed on the interpolated fragments during rendering. This is done before the fragment is modified by the tube blending function, so the tube replacement function value should reflect the scale and bias transformation.

In more complex scenarios, where a simplified branch contains children of its own, these children may need to be scaled again to effect their own simplifications. In Fig. 6, the topological zone labeled \( C_2 \) first is scaled about the saddle value \( S_2 \) when it is simplified. It is scaled again about saddle value \( S_1 \) when \( C_1 \) (the parent of \( C_2 \)) is simplified. Since a topological zone can have many parents, we compose all of the scaling transformations into a single scale-and-bias pair. When a user prunes a branch, the resulting scaling transformation \( s \) is applied to all of the branch’s children by multiplying each child’s scale-bias pair by \( s \). Similarly, when the pruning operation is reversed, the inverse transformation \( s^{-1} \) is applied to all of \( b \)’s children.

8 Results

Fig. 7 shows the “fuel” data set, a \( 64 \times 64 \times 64 \) voxel data set resulting from a simulation of fuel injected into a combustion chamber, see http://www.volvis.org/. Using a contour tree-based segmentation, it is possible to reveal internal structures (shown in color) while rendering surrounding structures using a low opacity in gray scale to provide context. Before interaction, the branch decomposition of the contour tree was simplified to 26 branches based on hypervolume. It is shown next to the rendered image to clarify the assignment of transfer functions to branches. Oversampling was used to obtain high-quality approximations of the gray-level isosurfaces that provide context.

Fig. 8a shows the “nucleon” data set, a \( 41 \times 41 \times 41 \) voxel data set resulting from a simulation of the probability distribution of a nucleon in a \(^{16}\text{O} \) nucleus, see http://www.volvis.org/. In addition to simplifying the branch decomposition to five branches for interaction, an extraneous branch was manually pruned. Here, we
used the topologically-based segmentation to simulate one of the topological attributes in Takeshima et al.’s work. We modified the opacity manually based on inclusion level to improve visibility of internal structures. Fig. 8b shows the same data set with a transfer function that emphasizes isosurfaces less and shows more of the internal structure of the nucleon. We emphasize regions around two minima in the “interior” while deemphasizing the “exterior” by choosing a much higher transparency and less saturated colors. Without the use of a global transfer function, it would not be possible to highlight these internal structures without having them occluded by external regions characterized by the same value range.

Figs. 9 and 10 show the results of applying our method of topology simplification to the fuel data set. In Fig. 9, we see an outside, isosurface-like view showing that the two separate components, corresponding to the “crown” and “shaft,” have been merged. The appearance of the crown is changed somewhat—that topological zone has been scaled down to allow for a monotone path to be created between the two components.

In Fig. 10, the same data set is shown by means of a single slice through the center of the data, illustrating the evolution of the interior values of the data set. In this figure, the unsimplified data is shown at the top, followed by versions simplified with flattening [20], gradient reversal [35], and our new method. As we can see, the effect of flattening in Fig. 10b is to replace a region of varying isovales with a single isovale. In contrast, gradient reversal in Fig. 10c bridges the saddle, allowing some of the higher isovales to remain. However, it is apparent both in the connection of crown to shaft at the right hand end, and the interior development at the left-hand end, that this...
narrow connection does a poor job of approximating the overall field. Finally, Fig. 10d shows the effect of the swept tubes, which broaden the connections to produce a smoother connection between components.

The value of this simplification method can also be observed in the movie accompanying this paper, which shows the evolution of the fuel data set using volume renderings for which a thin ramp function is varied over the isovalue range. In this movie, the data is shown in three forms: unsimplified, simplified by gradient reversal, and simplified with tube sweeping. As we see, the tube sweeping is a better approximation of the scalar field than the gradient reversal, exactly as predicted. Furthermore, simplified components gradually grow instead of popping up immediately, leading to smoother transitions and giving a user a better understanding of connectivity between components.

Fig. 11a shows a normal chest CT scan, available at http://radiology.uiowa.edu/downloads. It consists of $384 \times 384 \times 240$ voxels and shows how our segmentation performs on a moderately sized real-world data set. The tree, shown in Fig. 11b, has been simplified using hypervolume as underlying priority measure. Two ribs are highlighted in green. Note that the segmentation does not capture the entire rib; it shows only the regions of high density which are isolated from the spine and sternum. In Fig. 11d, the ribs have been made transparent in order to emphasize the lungs. The low-density regions near the lung-air interface, shown in blue, are easily separated from those near the skin-air interface, shown in gray.

Fig. 12 summarizes memory utilization and frame rates for a hardware implementation of our method. The “aneurism” data set, see http://www.volvis.org/, is included as a medium-size example. For small data sets, such as the “fuel” data set, interactive frame rates are possible. However, frame rates for larger data sets are still sufficiently high for interactive transfer function

<table>
<thead>
<tr>
<th>Data set</th>
<th>Voxels</th>
<th>Pre-processing</th>
<th>Branches</th>
<th>Transfer function texture</th>
<th>Branch texture</th>
<th>Frame rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel</td>
<td>$64 \times 64 \times 64$</td>
<td>2 sec.</td>
<td>86</td>
<td>$256 \times 8 \times 32$ bits</td>
<td>$256 \times 1 \times 80$ bits</td>
<td>21 fps</td>
</tr>
<tr>
<td>Aneurism</td>
<td>$256 \times 256 \times 256$</td>
<td>2 min., 6 sec.</td>
<td>31,850</td>
<td>$1024 \times 1024 \times 32$ bits</td>
<td>$256 \times 128 \times 80$ bits</td>
<td>2.5 fps</td>
</tr>
<tr>
<td>Chest CT</td>
<td>$384 \times 384 \times 240$</td>
<td>42 min., 45 sec.</td>
<td>99,406</td>
<td>$1024 \times 512 \times 32$ bits</td>
<td>$512 \times 256 \times 80$ bits</td>
<td>1.3 fps</td>
</tr>
</tbody>
</table>

Fig. 11. Normal chest CT data set consisting of $384 \times 384 \times 240$ voxels. In (a), two ribs are highlighted in green; the lung is shown in violet. The tree has been simplified to 28 branches based on hypervolume. (b) shows the flat tree layout, and (c) shows the orrery layout. In (d), the ribs have been removed and the lung is emphasized.
specification. Rendering measurements are based on a 1.8 gigahertz AMD Athlon 64 system with one gigabyte of RAM and an NVIDIA 7800 graphics board with 256 megabytes of graphics RAM. Rendered images consist of $512 \times 512$ pixels. Due to memory requirements of slightly more than one gigabyte of RAM for the “chest CT” data set, preprocessing measurements were performed on a three gigahertz Xeon system with four gigabytes of RAM.

9 CONCLUSIONS AND FUTURE WORK

We have shown how to extend topological segmentation of the domain of trivariate scalar fields to direct volume rendering; how to define and edit spatially local transfer functions based on the topological structure encapsulated in the contour tree; and how to implement ray casting for topological zones. We believe that this method of topologically defining a segmentation greatly increases the flexibility of transfer function design and significantly extends the utility of direct volume rendering.

We plan to assess the impact of different contour tree simplification measures on resulting visualizations. Future work will also be directed at interface design. We plan to consider different visual representations of the contour tree, such as representations that show nesting properties of contours [42], [43] and also to incorporate genus changes [19] in the user interface. While adding these saddles to the augmented contour tree does not change the segmentation as no branching occurs, they can still provide additional valuable information.

In addition to assigning independent transfer functions to topological zones, we would use different material properties, or even entirely different rendering modalities as discussed in [7], [8]. It would be interesting to combine our method with some of the user interface elements described by Takahashi et al. [28] and Kniss et al. [4]. In particular, the ability to drag topological zones to different positions and generate an exposed view would be a powerful user interface component.

ACKNOWLEDGMENTS

This work was supported by the US National Science Foundation under contract ACI 9624034 (CAREER Award), through the Large Scientific and Software Data Set Visualization (LSSDSV) program under contract ACI 9982251, through the National Partnership for Advanced Computational Infrastructure (NPACI) and a large Information Technology (LSSDSV) program under contract ACI 9982251, through the Large Scientific and Software Data Set Visualization Hardware, “IEEE Visualization Conf. ’03, pp. 301-308, 2003.


Gunther H. Weber received the PhD degree in computer science from the University of Kaiserslautern in 2003. His research on visualization of adaptive mesh refinement data and topology-based methods for the exploration of volume data was performed in close collaboration with the Institute for Data Analysis and Visualization (IDA) at the University of California, Davis, and the Lawrence Berkeley National Laboratory (LBNL). Currently, he is a postdoctoral researcher at IDAV and the Berkeley Drosophila Transcription Network Project at LBNL. He is a member of the IEEE Computer Society.

Scott E. Dillard received the BS degree in computer science from the University of California, Davis (UC Davis), in 2004. He is currently a PhD student at UC Davis, working with the Institute for Data Analysis and Visualization (IDA). His research interests include graphics, visualization, image processing, and computational geometry.

Hamish Carr completed the PhD degree at the University of British Columbia in May 2004 and was appointed as a lecturer at the University College Dublin, effective September 2004. His research interests center around the application of topological analysis to computer graphics and to scientific and medical visualization, but more broadly include computer graphics, computational geometry, and geometric applications. He is a member of the IEEE and the IEEE Computer Society.

Valerio Pascucci received the PhD degree in computer science from Purdue University in May 2000 and the EE Laurea (master’s), from the University “La Sapienza” in Rome in December 1993, as a member of the Geometric Computing Group. He has been a computer scientist and project leader at the Lawrence Livermore National Laboratory, Center for Applied Scientific Computation (CASCS) since May 2000. Prior to his CASCS tenure, he was a senior research associate at the University of Texas at Austin, Center for Computational Visualization, CS and TICAM Departments. He is a member of the IEEE and the IEEE Computer Society.

Bernd Hamann received the PhD degree in computer science from Arizona State University in 1991. He serves as an Associate Vice Chancellor for Research and is a professor of computer science at the University of California, Davis. His main interests are visualization, geometric modeling, and computer graphics. He was awarded a 1992 US National Science Foundation (NSF) Research Initiation Award, a 1996 NSF CAREER Award, and a 2006 University of California Presidential Chair in Undergraduate Education. He is a member of the IEEE and the IEEE Computer Society.

> For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.