



## On the search of optimal reconstruction resolution

Erald Vuçini<sup>a,b,\*</sup>, Walter G. Kropatsch<sup>a</sup>

<sup>a</sup> Institute of Computer Graphics and Algorithms, Favoritenstrasse 9, 1863, 1040 Vienna, Austria

<sup>b</sup> VRVis Center for Virtual Reality and Visualization Research, Donau-City-Strasse 1, 1220 Vienna, Austria

### ARTICLE INFO

#### Article history:

Available online 20 October 2011

#### Keywords:

Topology  
Non-uniform representation  
Reconstruction  
Homology

### ABSTRACT

In this paper we present a novel algorithm to optimize the reconstruction from non-uniform point sets. We introduce a statistically-derived topology-controller for selecting the reconstruction resolution of a given non-uniform point set. Deriving information from homology-based statistics, our topology-controller ensures a stable and sound basis for the analysis process. By analyzing our topology-controller, we select an optimal reconstruction resolution which ensures both low reconstruction errors and a topological stability of the underlying signal. Our approach offers a valuable method for the evaluation of the reconstruction process without the need of visual inspection of the reconstructed datasets. By means of qualitative results we show how our proposed topology statistics provides complementary information in the enhancement of existing reconstruction pipelines in visualization.

© 2011 Elsevier B.V. All rights reserved.

### 1. Introduction

In the last decades, unprecedented technological growth and development have contributed to the overall improvement of the visualization pipeline, in particular for the processes of data acquisition and data enhancement. The traditional sources of volumetric data are simulations as well as data acquisition devices. The majority of these devices acquire data on uniform (Cartesian) lattices. In an effort to study larger and more complex problems, there has been a move towards non-uniform (irregular) data representations, since they offer a way of adapting the measure location (or sample points) according to the importance (variance) of the data.

While the acquisition of data on non-uniform grids has become wide-spread, the available tools for processing, filtering, analysis, and rendering of data are most efficient for uniform representations. One of the methods of dealing with non-uniform data is to convert them into efficient uniform representations and apply standard tools. The key issue in such conversion (reconstruction) is the selection of the proper resolution of reconstruction. The non-triviality of this problem lies on the fact that the ground-truth signal is unknown. Fig. 1 shows a 2D example of such a problem. In Fig. 1(a) we show the original image. Fig. 1(b)–(d) show the uniform representations after doing reconstruction with three different resolutions. The reconstructed signal is represented by the uniform grid points, whereas values at the interior of each cell can be obtained through interpolation, e.g., nearest neighbor, linear or cubic interpolation. The accuracy of interpolation depends

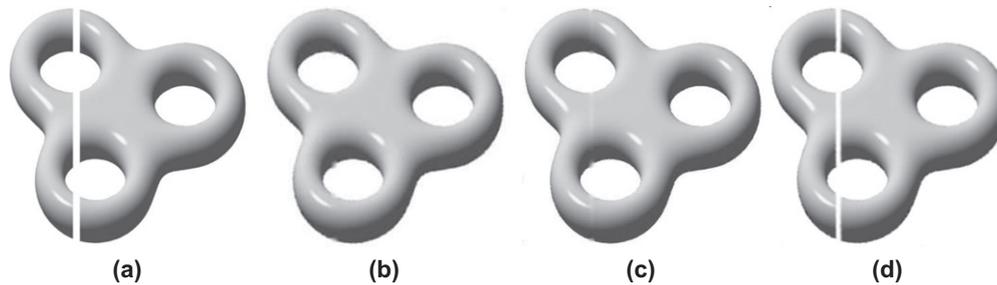
heavily on the resolution of the uniform representations. A coarse resolution of reconstruction requires less memory consumption but will smooth the signal in areas with sharp transitions, resulting in visual artifacts or high errors. A too fine resolution of reconstruction will introduce memory overheads while achieving a better accuracy. A trade-off between accuracy and memory efficiency is required.

On the other side, since the original signal is unknown, it is difficult to evaluate the quality of reconstruction. Several methods require signal bandlimitness and shift-invariance as necessary preconditions for guaranteeing a perfect and stable reconstruction. Both these constraints generally do not hold for signals we want to analyze and process in the visualization domain. A commonly followed approach for assessing the quality of reconstruction is to measure the reconstruction error, e.g., the root mean square error (RMSE). In addition to the fact that the RMSE can be misleading due to his averaging behavior, in many scenarios low RMSEs may still result in artifacts in the reconstructed data. Hence, direct visual inspection is required, introducing the need of more time overhead and user interaction. In Fig. 1(b) and (c), we see such artifacts as topological changes. Although the function is well-approximated in areas where non-uniform points are given, the same does not hold for the empty areas (with no non-uniform points). Since the RMSE measures the error only in the input non-uniform points, there is no possibility to realize the formation of such artifacts. This scenario can happen even when the RMSE is low (in comparison to some user-defined threshold).

The selection of the optimal reconstruction resolution will be the central question we are trying to answer in this paper. We assume that we can only afford a single resolution and we make suggestions on how this resolution can be best obtained. This assumption is

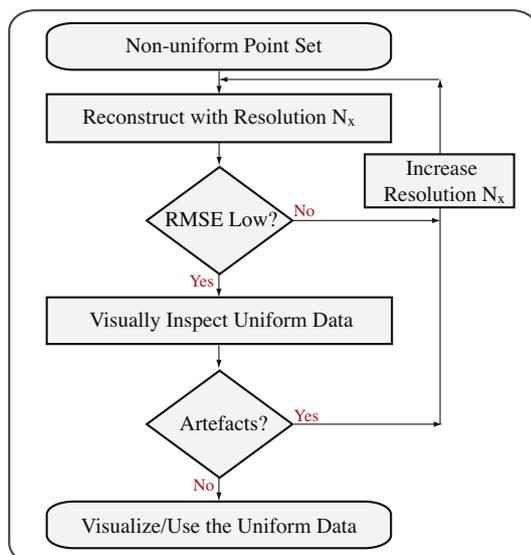
\* Corresponding author at: VRVis Center for Virtual Reality and Visualization Research, Donau-City-Strasse 1, 1220 Vienna, Austria.

E-mail address: [vucini@vrvis.at](mailto:vucini@vrvis.at) (E. Vuçini).



**Fig. 1.** 2D example of the view of a (cut) torus: (a) original image of resolution  $512 \times 512$ , (b) reconstruction with resolution  $256 \times 256$ , (c) reconstruction with resolution  $358 \times 358$ , and (d) reconstruction with resolution  $512 \times 512$ . The input points for the reconstructions are chosen randomly from the original image (we select 30% of the total number of points, i.e., pixels). We see in Fig. (b) and (c) that the gap between the two parts of the torus has disappeared, yielding an unwanted result. Increasing the resolution of reconstruction improves the reconstruction quality and minimizes the artifacts (Fig. (d)).

applicable for non-uniform data, where the distribution of samples is even (in the sense of a discrepancy measure), e.g., ultrasound data, seismic data or missing samples data from communication theory. We improve the efficiency of the reconstruction workflow by supplying statistical information from homology analysis. Unlike data frequency histograms, which show data value related statistics, our proposed topology controller gives important complementary information of the data topological changes. In a traditional reconstruction framework, we would follow these steps when trying to reconstruct a non-uniform points set  $\mathbf{P}$ : (1) reconstruct  $\mathbf{P}$  with different resolutions, (2) compute the RMSEs for each reconstruction resolution, (3) select those reconstructions that have RMSEs lower than a user defined threshold, and (4) visually inspect the reconstructed data for possible artifacts, e.g., by means of volume rendering. A schematic view of this reconstruction pipeline is displayed in Fig. 2. With our topology-controller we eliminate the time consuming step of visual inspection. Instead, by means of a simple graphical plot, we are able to suggest a resolution of reconstruction which is optimal w.r.t. our topology-controller. Another issue which negatively effects the traditional pipeline is the difficulty or impossibility to visualize very large datasets due to graphics hardware constraints. Although the reconstruction process is possible, the visualization of such data is not trivial. Hence, methods that reconstruct, analyze and verify the quality in an offline-preprocessing mode are a viable option.



**Fig. 2.** Schematic view of a traditional pipeline for the reconstruction of non-uniform point sets to uniform representations when the target resolution is unknown.

The topological information together with error measurements improves the quality assessment of the reconstruction and reduces the need of visual inspection. In our example, we will be able to clearly discriminate between the cases shown in Fig. 1(b)–(d). The latest one has a stable topological behavior with regard to our proposed topology-controller. It is worth to mention, that in our framework we do not consider the geometrical changes in the data. However, such cases are handled by the error measurement. After a specific resolution, the errors stabilize, and further increase of resolution does not produce different geometrical results. The reconstruction framework would yield only oscillations of the signal as merging features (illustrated in Fig. 1), or the creation/destruction of topological features.

In Section 2 we give a summary of the existing research related to our work. We introduce our topology controller in Section 3, and explain the main modules of the proposed iterative algorithm. In Section 4 we show results w.r.t. the usage of the topology-controller and assess the value of its applicability. The tuning of the algorithm parameters is discussed in Section 4.2. Finally, conclusions and ideas for future work are summarized in Section 5.

## 2. Related work

Our proposed method connects concepts starting from signal reconstruction to topology, statistics and their usage in the visualization domain.

Non-uniform data reconstruction (approximation) is a recent, fast growing research area. A number of approaches reconstruct non-uniformly sampled data, especially for one- and two-dimensional signals. Most of the methods are based on the reconstruction of the data by solving large systems of equations (Feichtinger et al., 1995; Grishin and Strohmer, 2004). Nielson (1993) presented an overview of several approximation techniques for non-uniform point sets. While each technique performs best only in particular cases, the use of local compact operators is considered the fastest approach. Perhaps the most popular approach for approximating non-uniform data is based on Radial Basis Functions (RBFs). They have been used in surface (Ohtake et al., 2004) as well as volumetric (Jang et al., 2006) approximation and reconstruction techniques.

Arigovindan et al. (2005) proposed to use B-splines in a multi-grid framework for the reconstruction of non-uniform 2D data. Vućini et al. (2008, 2009) extended these ideas to 3D volumes and large datasets. B-splines, with their smoothness and compact support, offer optimal conditions for fast and accurate reconstruction results. They are related to RBF-based approaches since B-splines are very good approximators of thin-plate splines, which in turn are widely used RBFs in approximation theory. In this paper we build on these ideas.

Most of the above-mentioned approaches consider the resolution of reconstruction as known *a priori*. In order to find the resolution allowing exact reconstruction a lower bound on the minimal distance between two sampling positions has to be assured. For general shift-invariant spaces a Beurling density  $D \geq 1$  is necessary for a stable and perfect reconstruction (Aldroubi and Gröchenig, 2000). In topology analysis, in order to be able to provide a topological-stable reconstruction, the object (signal) taken into consideration has to be  $r$ -regular. The related literature is mainly related to the problem of surface reconstruction (Stelldinger et al., 2007; Meine et al., 2009). Recently, Vućini et al. (2009) proposed the usage of the  $\sigma_{avg}$  concept for selecting an optimal resolution of reconstruction. While this approach works well in the proposed reconstruction pipeline, still the method is based on heuristically derived assumptions and no clear proof is given that this is an optimal characteristic that works with other reconstruction pipelines.

Topology is gaining more and more significance in the analysis of multi-dimensional data, since it offers additional insight in the effort of data understanding. Due to the complexity of the data, techniques for providing a simplified view are required. Topology analysis has been successfully linked to fields related to isosurface selection (Bajaj et al., 1997; Takahashi et al., 2004), topological downsampling and simplification (Kraus and Ertl, 2001; Gyulassy et al., 2005; Natarajan and Pascucci, 2005), topology-guided analysis and navigation in scalar and time varying data (Weber et al., 2007; Bremer et al., 2010), and feature tracking and evolution (Silver and Wang, 1998; Weber et al., 2009). Carr et al. (2010) have presented a generalized framework consolidating the theory and application of the contour spectrum concept. Most of these works are based on Morse theory, in particular the Morse–Smale complex (Milnor, 1963), and have concentrated in reporting topological information related to 0-dimensional homology, i.e., connected components. Our approach derives homological statistics by computing the Betti numbers, which in turn are connected to Morse theory since they are upper-bounded by the respective number of critical points (Matsumoto, 2001). Topological and persistence information has been used also for shape comparison and feature classification (Cerri et al., 2007; Chazal et al., 2009). Recently, Bendich et al. (2010) have introduced the persistence diagram, an analysis tool for displaying topological information of higher dimensions. Although their proposed method is very novel and promising, it is constrained to data given as simplicial complexes and suffers from very high memory requirements.

All the above-mentioned methods provide extensive information, which is difficult to interpret without the appropriate statistical analysis. In our method we use statistical topological information for guiding the selection of resolution of reconstruction. As a result, our interface gives important cues that reduce the necessity of human's visual inspection of the data.

### 3. Topology-based analysis

We propose an algorithm with two main modules: (1) the variational reconstruction module, and (2) the module that derives the statistical homology information. Both modules are integrated in the main iterative procedure which extracts useful statistical information related to the reconstruction process. Through this information we will be able to select a resolution of reconstruction that has both a low RMSE and topological stability with regards to our defined topology-controller ( $\mathcal{TC}$ ).

#### 3.1. Variational reconstruction basics

Variational reconstruction is a well-known technique applied to solving ill-posed problems such as the reconstruction from

non-uniform point sets. The variational functional is formulated so that it provides a solution close to the input points, while regularizing the smoothness in order to prevent discontinuities.

Given a set  $\mathbf{P}$  of sample points,  $\mathbf{p}_i = (x_i, y_i, z_i)$ ,  $i = 1, 2, \dots, M$ , let  $f_i$  be the scalar values associated with  $\mathbf{p}_i$ . We define the B-spline approximation through the form:

$$F(x, y, z) = \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} \sum_{m=0}^{N_z-1} c_{k,l,m} \beta^3(x-k) \beta^3(y-l) \beta^3(z-m) \quad (1)$$

where  $\beta^3(x)$  is the cubic B-spline basis function,  $c_{k,l,m}$  are the B-spline coefficients and  $(N_x, N_y, N_z)$  is the resolution of the axis-aligned bounding box of our non-uniform dataset. Cubic B-splines do not enjoy the interpolation property, but with real-world data where noise is always present, approximative (not-interpolating) splines are better suited for the reconstruction process. Furthermore, they provide the best quality for a given computational cost (Thévenaz et al., 2000).

In order to determine the B-spline coefficients the following cost function is minimized:

$$C(F) = \sum_{i=0}^{M-1} \|F(\mathbf{p}) - f_i\|^2 + \lambda \iiint \|D^p F\|^2 dx dy dz \quad (2)$$

where  $\lambda$  is a parameter that controls the smoothness and the second term is the regularization functional that uses Duchon's seminorms  $D^p F$  (Duchon, 1979).

The key idea of the variational reconstruction is to build a linear system related to the unknown B-spline coefficients  $c_k$  and to solve it by minimizing the cost function. Once we solve the equation, we have  $N_x \times N_y \times N_z$  B-spline coefficients defined at the reconstruction-grid positions. We can compute  $F(\mathbf{p})$  (a  $C^2$ -continuous function) at any position  $\mathbf{p} \in V$ , where  $V$  is the volume enclosing the bounding box of the non-uniform point set. For a deeper insight into the method we refer the reader to Vućini et al. (2009).

#### 3.2. Statistical homology

Betti numbers are topological features, proved to be invariants with regard to continuous deformation, and used to extend the polyhedral formula to higher dimensional spaces (Munkres, 1984). The first Betti numbers  $b_0$ ,  $b_1$  and  $b_2$  have the following intuitive definitions:  $b_0$  is the number of connected components,  $b_1$  is the number of handles (tunnels),  $b_2$  is the number of three-dimensional holes (voids).

Betti numbers are characteristics of point set complexes, given either as simplicial or cubical complexes. They depend on the point coordinates and are not affected by the scalar values associated with the points. Typically they are studied in relation with isosurfaces (level sets), but such usage introduces unstable behavior due to the significant noise, present in the data, especially in real-world datasets. To have a better understanding on which isovalues play more significant role in the birth/death of homological structures (i.e., components, tunnels, voids), we use the concept of superlevel sets. Given a volume  $V$  as a set of voxels  $v$  with values specified by the function  $F:V \rightarrow R$ , we define the superlevel set  $S_V^+$  as:

$$S_V^+(t) = \{v \in V | F(v) \geq t\} \quad (3)$$

This structure is less influenced by noise and will allow us to study the stability of Betti numbers over different data-ranges (in the sense of persistence Edelsbrunner et al., 2002). We compute the Betti numbers for cubical cells given by the superlevel set  $S_V^+(t)$ . We deal with 3D data, hence we estimate only the first three Betti numbers. We define as  $\mathcal{H}_V^\tau(t)$  the function that computes the  $\tau$ -dimensional Betti number of the superlevel set  $S_V^+(t)$ , for  $\tau = 0, 1, 2$ .

### 3.3. Iterative Topology-based Algorithm (ITA)

Our iterative algorithm takes as input a non-uniform point set ( $\mathbf{P}$ ), a minimum ( $N_{min}$ ) and maximum ( $N_{max}$ ) resolution of reconstruction, and the number  $L$  of superlevel sets we build for each reconstructed volume (see Algorithm 1).

---

**Algorithm 1.** ITA ( $\mathbf{P}, N_{min}, N_{max}, \Delta, L$ )
 

---

- 1: **for**  $i = N_{min}$  to  $N_{max}$  **do**
  - 2: determine  $V_i$  by solving variational reconstruction for  $\mathbf{P}$  on resolution  $i$
  - 3: build  $L + 1$  superlevel sets  $S_V^+(t_l)$  ( $l = 0, 1, \dots, L$ ) where:  
 $t_l = Max_{value} \cdot (1 - l/L)$
  - 4: **for all**  $(L + 1)$  superlevel sets  $S_V^+(t_l)$  **do**
  - 5: compute  $\mathcal{H}_{V_i}^0(t_l)$ ,  $\mathcal{H}_{V_i}^1(t_l)$  and  $\mathcal{H}_{V_i}^2(t_l)$
  - 6: **end for**
  - 7: build  $H_{V_i}^0$ ,  $H_{V_i}^1$  and  $H_{V_i}^2$
  - 8:  $i = i + \Delta$
  - 9: **end for**
  - 10: compute the topology controller ( $\mathcal{TC}$ )
- 

The iterative algorithm starts with determining the volume  $V_i$  from the B-spline coefficients estimated from the variational reconstruction of the non-uniform point set  $\mathbf{P}$  (line 2). The resolution of reconstruction  $N_x \times N_y \times N_z$  is specified by the loop-variable  $i$  (loop-variable  $i$  is augmented in each step by the variable  $\Delta$ ). By varying  $N_x$ ,  $N_y$  and  $N_z$  are determined automatically by the proper aspect ratio of the axis-aligned bounding box enclosing the given non-uniform data points. We build  $L + 1$  superlevel sets for each volume  $V_i$  (line 3). The number of superlevel sets depends on the desired accuracy. Generally we set  $L = Max_{value}$ , where  $Max_{value}$  is the maximum value for the underlying data format, e.g.,  $Max_{value} = 255$ . For each superlevel set we compute the respective Betti numbers represented by  $\mathcal{H}_{V_i}^0(t_l)$ ,  $\mathcal{H}_{V_i}^1(t_l)$  and  $\mathcal{H}_{V_i}^2(t_l)$  (line 4–6). We denote as  $H_{V_i}^\tau$  the row-vector composed of the  $(L + 1)\mathcal{H}_{V_i}^\tau(t_l)$  values of superlevel sets for each  $\tau$ -dimensional Betti number (line 7).  $H_{V_i}^\tau$  can be considered as the  $\tau$ -dimensional homological signature of the point set  $\mathbf{P}$  in resolution  $i$ .

After computing all homological statistics for each superlevel set of each resolution, we can define the topology controller  $\mathcal{TC}(i)$  as:

$$\mathcal{TC}(i) = \frac{1}{\alpha_0 + \alpha_1 + \alpha_2} \sum_{\tau=0}^2 \alpha_\tau \cdot \frac{\|H_{V_i}^\tau - H_{V_{N_{max}}}^\tau\|}{\|H_{V_{N_{max}}}^\tau\|} \quad (4)$$

where the weights (coefficients)  $\alpha$  control the impact of the respective  $\tau$ -dimensional homological statistics ( $H^\tau$ ) on the topology controller. In simpler words,  $\mathcal{TC}(i)$  computes the relative error of  $H_{V_i}^\tau$  with regard to  $H_{V_{N_{max}}}^\tau$ , which is the homological signature of the point set  $\mathbf{P}$  in the maximum resolution. The main motivation for using the relative error measurement is to achieve comparable values of the topology controller for different datasets. The comparison with the homological signature of the maximum resolution is also most meaningful, since by that resolution we expect that all topological artifacts have been minimized. The division by  $(\alpha_0 + \alpha_1 + \alpha_2)$  in Eq. (4) ensures the normalization of the topology-controller. In case, when  $H_{V_{N_{max}}}^\tau = 0$ , we consider the respective term as trivial topology and discard it.

## 4. Implementation and results

The variational reconstruction is designed and implemented by solving a minimization problem through multi-grid iterations. In

our framework we have used the settings as described by Vućini et al. (2009). The extraction of the homological information related to the Betti numbers is performed in cubical cells, i.e. uniform volumes, as described by Kaczynski et al. (2004). Due to recently developed reduction homology algorithms, these homology computations can be performed very fast.<sup>1</sup> Both the variational reconstruction and the homology computation modules are implemented as an offline preprocessing step and their complexity and computational times are the same as reported by the respective authors.

We use the topology-controller for evaluating the quality of reconstruction. We consider a resolution of reconstruction as optimal if it fulfills two criteria: (1) low reconstruction errors (RMSE), and (2) minimal number of artifacts. Since artifacts are mainly described as topological changes/deformations, we use our topology-controller to quantify these changes. We analyze graphical plots of  $\mathcal{TC}$  with regard to a changing resolution and we set default weight values ( $\alpha_0 = \alpha_1 = \alpha_2 = 1$ ) in Eq. (4). We attach to these plots also the graphs of  $\mathcal{TC}_\tau$ , which measure the  $\tau$ -dimensional homological statistics.  $\mathcal{TC}_\tau$  is derived from  $\mathcal{TC}$  by setting the respective  $\alpha_\tau$  equal to one and the other two weights equal to zero. We postulate that a resolution of reconstruction  $N_x$  is optimal with regard to the proposed topology-controller if  $\mathcal{TC}(N) \leq 0.2$ , for every  $N$  such that  $N \geq N_x$  (see Section 4.2 for more details). In the graphical plots of the topology-controller we have omitted the plotting of  $\mathcal{TC}$  for  $N = N_{max}$  since  $\mathcal{TC}(N_{max} = 0)$ .

We tested our framework on non-uniform datasets acquired from different sources. For original non-uniform point sets there is no possibility to exactly measure the accuracy of a reconstruction or visualization method at positions not known *a priori*, since there is no ground truth. In order to tackle such issue and to better understand the behavior of our proposed topology-controller we created non-uniformly sampled data from regular datasets by adaptively sampling them. For the adaptive sampling of the data we used a 3D Laplacian kernel.<sup>2</sup> After convolving the data with this 3D kernel we sorted the point values according to their magnitudes and retain only those points that have the biggest absolute values, i.e., 20% of all points in our experiments. This is equivalent to keeping the points on both sides of boundary regions.

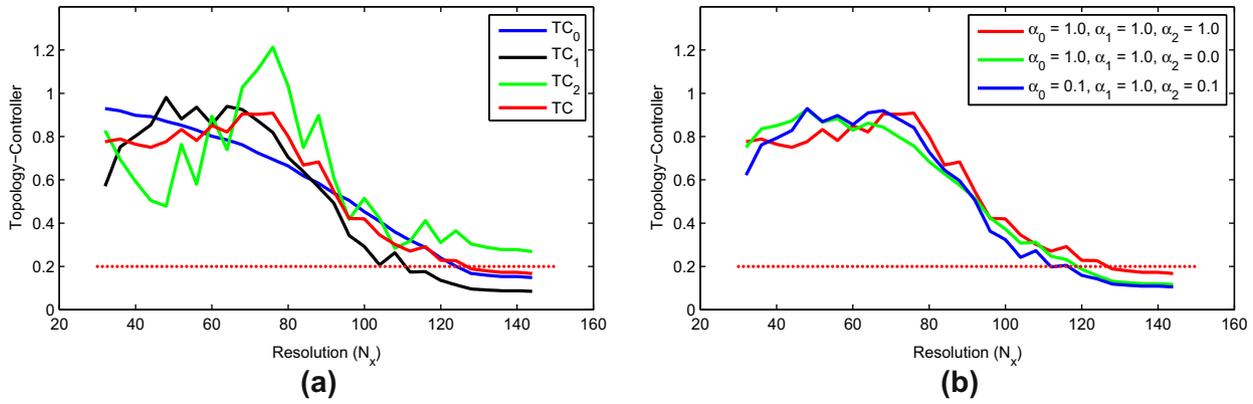
In Fig. 3, graphical plots of  $\mathcal{TC}$  for the **Aneurism Laplacian** dataset are given. The Aneurism Laplacian dataset consists of 419,430 non-uniform points extracted from a  $128 \times 128 \times 128$  uniform representation, and it is a rotational C-arm X-ray scan of the arteries of the right half of a human head. In Fig. 3(a) we see that our topology criteria is fulfilled when  $N_x \geq 125$ . This conclusion coincides with the original resolution of the Aneurism dataset. Fig. 3(b) shows further graphs of  $\mathcal{TC}$  with different weight settings. These graphs are especially useful when we have *a priori* knowledge of the nature of the data we are working with. In the case of the Aneurism dataset, the most important elements are the connected components and tunnels. Hence, setting the weights  $\alpha_0 = \alpha_1 = 1$  and  $\alpha_2 = 0$  allow us to better track and quantify important topological changes (see green line in Fig. 3(b)).

In Fig. 4(a) we show a graphical plot of  $\mathcal{TC}$  related to an originally non-uniform dataset. The **Bypass** dataset is a simulation from a laminar-turbulent transition in a boundary layer that is subject to free stream turbulence. It consists of 7,929,856 non-uniform points in a curvilinear grid with non-uniform spacing along the y-axis. The visualization of this simulation is of great importance to better analyze how the “bypass” of the Tollmien–Schlichting (TS) waves develops (Schlatter, 2001). In Fig. 4(b) we show the graph of the RMSE as we increase the resolution of reconstruction.<sup>3</sup> While the

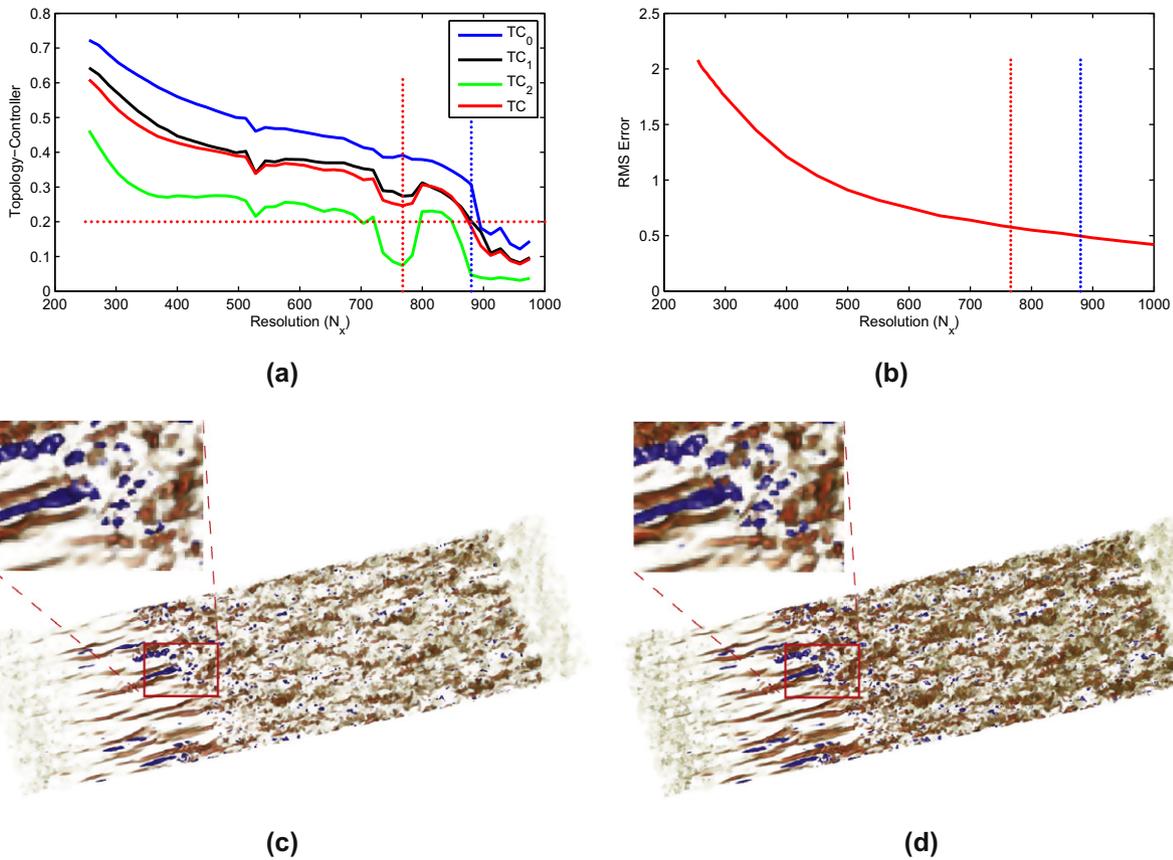
<sup>1</sup> See Computational Homology Project at <http://chomp.rutgers.edu>.

<sup>2</sup> See Appendix A for the 3D Laplacian kernel formulation.

<sup>3</sup> See Appendix B for RMSE formulation.



**Fig. 3.** Graphical plots of topology-controller for the Aneurism dataset. The graph with default setting of  $TC$ , ( $\alpha_0 = \alpha_1 = \alpha_2 = 1$ ), is shown with the red line in (a) and (b). The horizontal red cross-hair shows the topology-controller threshold. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 4.** (a) Graphical plot of topology-controller for the Bypass dataset. The graph with default setting of  $TC$ , ( $\alpha_0 = \alpha_1 = \alpha_2 = 1$ ), is shown with the red line. The horizontal red cross-hair shows the topology-controller threshold. (b) Graph showing the RMSE according to the changes of the resolution of reconstruction  $N_x \times N_y \times N_z$  for the Bypass dataset. The vertical red cross-hair shows the optimal resolution as suggested by Vućini et al. (2009), and the blue one shows the resolution suggested by the topology-controller. (c) Rendering with close-ups of the Bypass dataset for optimal reconstruction resolution (766 × 92 × 192), as depicted in (Vućini et al., 2009), RMSE is 0.61. (d) Rendering with close-ups of the Bypass dataset optimal reconstruction resolution (880 × 106 × 220), as depicted by our topology-controller, RMSE is 0.49. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

reconstruction error is below one, starting from resolution  $N_x \approx 500$ , there is no clear intuition what is the optimal resolution. Vućini et al. (2009) selected an optimal resolution of  $766 \times 92 \times 192$  based on their  $\sigma_{avg}$  assumption. From the analysis of Fig. 4(a) and (b) we agree that although there is a local minimum at  $N_x = 766$ , the topology criterion is achieved after  $N_x = 880$ . In Fig. 4(c) and (d) we show two renderings with close-ups of the Bypass dataset reconstructed with resolutions  $766 \times 92 \times 192$  and  $880 \times 106 \times 220$  respectively.

Visual differences are clear especially for the semi-transparent structures. The opaque structures (color-mapped to blue) differ in size, hence resulting in different superlevel sets statistics. This in turn is the main reason for the differences in the  $TC$  graphs.

In Table 1 we show results of our topology-controller related to the resolution selection for several non-uniform datasets. We see that the suggested resolutions by the proposed topology-controller vary  $\pm 15\%$  from the results presented by Vućini et al. (2009).

**Table 1**

Comparisons of reconstruction resolutions for different non-uniform datasets as proposed by Vućini et al. (2009) and the topology-controller.

Dataset		Vućini et al. (2009)		Topology-controller ( $\mathcal{TC}$ )	
Name	Points	Resolution	RMSE	Resolution	RMSE
Oil	29,094	$38 \times 40 \times 38$	0.19	$41 \times 44 \times 41$	0.16
Natural Convection	68,921	$61 \times 61 \times 61$	0.63	$55 \times 55 \times 55$	0.71
Synthetic Chirp	75,000	$64 \times 64 \times 64$	1.12	$58 \times 58 \times 58$	1.18
Bypass	7,929,856	$766 \times 92 \times 192$	0.61	$880 \times 106 \times 220$	0.49
Blunt-Fin	40,960	$93 \times 36 \times 25$	1.14	$102 \times 39 \times 28$	1.11

Although the values of reconstruction resolutions and RMSE are similar, we can now be more confident in the quality of reconstruction due to the topology controller.

#### 4.1. Timings and complexity

In Table 2 we show the performance of our algorithm for different datasets and settings. Parameter  $L$  decides how many superlevel set signatures will be computed. On the other side,  $N_{min}$ ,  $N_{max}$  and  $\Delta$  specify how many times the whole algorithm will run. For the Aneurism dataset the homology signature is computed for 32 data sets ( $\Delta = 4$ ), while for the Neghip and Natural Convection is 64 and 48, respectively ( $\Delta = 1$ ). The algorithm scales linearly with parameter  $L$ .

The run-time complexity of the overall algorithm depends on the complexity of the two main modules. We can roughly derive the complexity by the formula:  $\frac{N_{max}-N_{min}}{\Delta} \cdot (O_R + 3 \cdot (L+1) \cdot O_H)$ , where  $O_R$  and  $O_H$  are the run-time complexities of the variational reconstruction module and the homology statistics computing module, respectively. The complexity of the variational reconstruction module is linear with regard to the grid points ( $N_x \cdot N_y \cdot N_z$ ). The worst-case complexity of the homology module is cubic with regard to the input points, i.e., in our case the superlevel set points. However, similar to many other researchers, we have observed a linear run-time complexity for the homology module. Hence, we can state that the worst-case complexity of our algorithm is cubic, but the observed complexity is linear.

#### 4.2. Parameter tuning

Our iterative algorithm is based on the selection of several parameters which affect the general performance and output. These parameters are:  $N_{min}$ ,  $N_{max}$ ,  $\Delta$ ,  $L$ , the topology-controller threshold and the  $\alpha$ -weights.

##### 4.2.1. Selecting $N_{min}$ , $N_{max}$ , $\Delta$ and $L$

$N_{min}$  and  $N_{max}$  are tied to the number of points in the set  $\mathbf{P}$  and to the hardware constraints. One should not choose very high resolution  $N_{max}$ , such that it compromises analysis performance, yet, it should also not be too coarse  $N_{min}$  to avoid computing irrelevant topological information. While the selection of a too coarse  $N_{min}$  does not affect the output result of our algorithm, the selection of  $N_{max}$  is more crucial due to Eq. (4). In cases when the non-uniform points are evenly distributed across the volume, setting  $N_{max}$  such that the number of voxels in  $V$  is approximately five times the

number of the input non-uniform points, is a reasonable trade-off between computational burden and output result. Our experimental results show that increasing  $N_{max}$  beyond such bound does not affect the selection of reconstruction resolution.

The parameters  $\Delta$  and  $L$  mainly affect the performance of the algorithm.  $\Delta = 1$  means that we explicitly reconstruct the non-uniform points with each resolution in the range  $[N_{min}, N_{max}]$ . From our experimental results, we have concluded that while it is acceptable to set  $\Delta = 1$  for small range datasets, for larger datasets, e.g., the Bypass dataset, this would introduce very large computation times. As shown in Table 2 we have selected  $\Delta = 16$  for the Bypass dataset.

Parameter  $L$  controls the refinement of the homology signature. Based on our testings, we have concluded that selecting  $L = \text{Max}_{value}/5$ , gives detailed information about the signature, ensuring a balance with the computational effort.

##### 4.2.2. Selecting the $\mathcal{TC}$ -threshold

The topology-controller threshold is the parameter that mostly affects the usage of our algorithm. We tested the algorithm with several datasets coming from different acquisition sources. From our experimental observations, we have concluded that selecting a reconstruction resolution that fulfills the condition  $\mathcal{TC} \leq 0.2$ , leads to consistent and reliable results. However, the interactivity of the framework let us decide the resolution to depict, based on the topology-controller graph.

##### 4.2.3. Selecting the weights

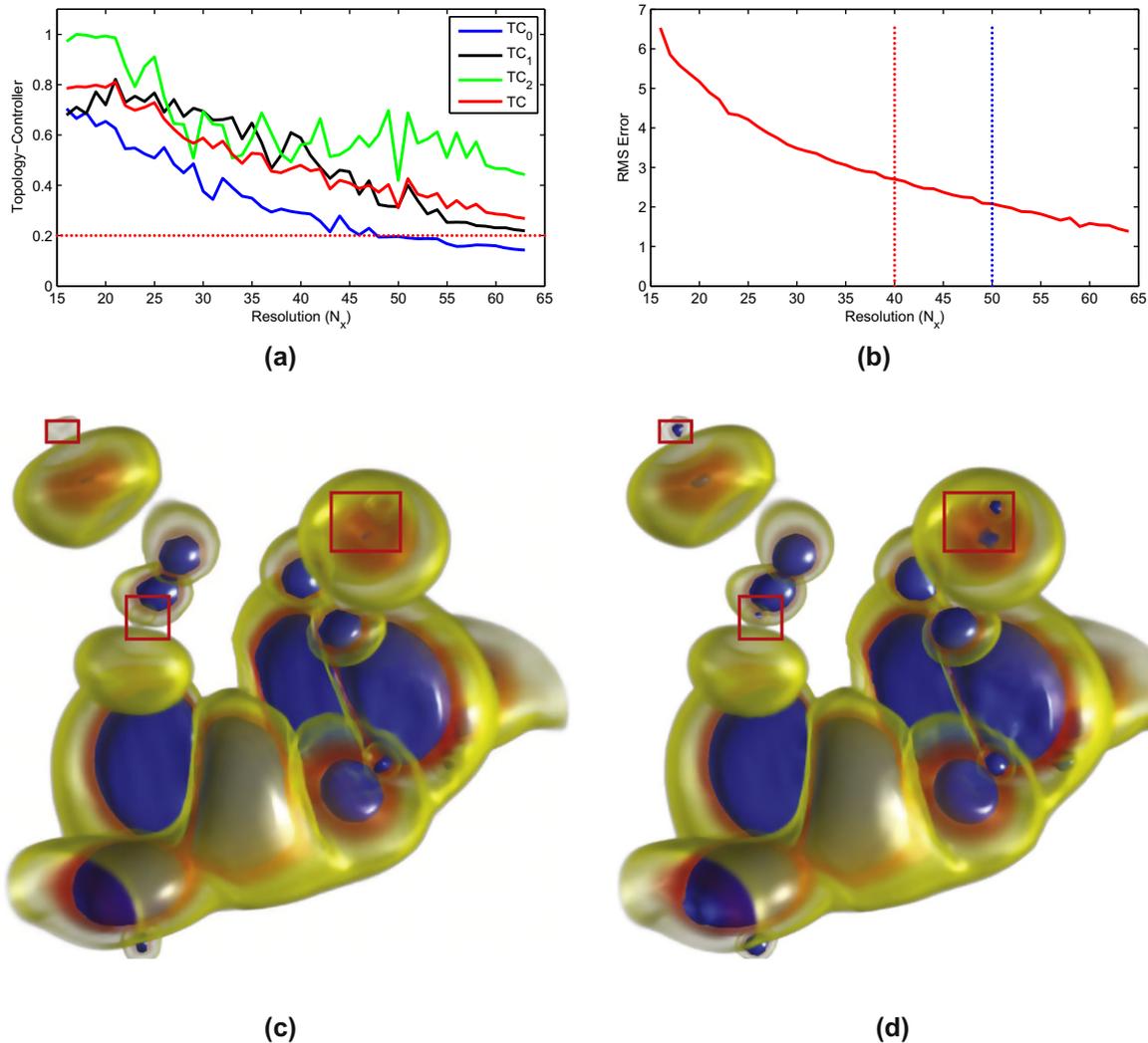
The selection of weights, can be driven by the mean value of the respective  $\tau$ -dimensional homological signature. High mean is mapped to a high weight value and vice versa. In this way, we decrease the influence of homological changes due to noise and artifacts present in the data. *A priori* knowledge of the data also helps in the selection of weights. E.g., if we know that the data mainly consists of handles (tunnels), we may select  $\alpha_1$  as the most prominent weight.

In a traditional reconstruction pipeline, the RMSE is an important criteria for evaluating the reconstruction quality. Generally speaking, RMSEs lower than one, lead to good reconstruction quality. For some data, the  $RMSE \leq 1$  condition is obtained only when a very high resolution is selected. This derives from the approximative behavior of the variational approach. In presence of noise, the variational approach tends to report higher RMSEs due to its smoothing effect. With the increase of resolution, no major changes happen in the data, hence a high resolution of reconstruc-

**Table 2**

Times for the computation of the topology-controller for different datasets and settings.

Dataset		Times (min)						
Name	Points	$N_{min}$	$N_{max}$	$\Delta$	$L = \text{Max}_{value}/10$	$L = \text{Max}_{value}/5$	$L = \text{Max}_{value}$	
Neghip	52,428	16	80	1	6.47	13.82	60.45	
Natural Convection	68,921	16	64	1	10.98	22.03	107.13	
Aneurism	419,430	32	160	4	4.15	8.35	40.77	
Bypass	7,929,856	256	1024	16	668.75	1390.93	6712.80	



**Fig. 5.** (a) Graphical plot of topology-controller for the Neghip dataset. The graph with default setting of  $\mathcal{TC}$ , ( $\alpha_0 = \alpha_1 = \alpha_2 = 1$ ), is shown with the red line. The horizontal red cross-hair shows the topology-controller threshold. (b) Graph showing the RMSE according to the changes of the resolution of reconstruction  $N_x \times N_y \times N_z$  for the Neghip dataset. The vertical red cross-hair shows resolution  $N_x = 40$ , and the blue one shows the resolution suggested by the topology-controller,  $N_x = 50$ . (c) and (d) Renderings with highlighting of the Neghip dataset with: (c) reconstruction resolution  $40 \times 40 \times 40$ , RMSE is 2.63 and (d) optimal resolution  $50 \times 50 \times 50$ , as depicted by our topology-controller, RMSE is 2.05. The three areas with greater changes are highlighted with red-border rectangles. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

tion results in inefficient use of memory and storage resources. Our topology-controller offers an intuitive solution to such problem, since it measures the relative topological difference to a specified target resolution ( $N_{max}$ ).

Fig. 5 shows result related to the **Neghip Laplacian** dataset. This dataset consists of 52,428 non-uniform points extracted from a  $64 \times 64 \times 64$  uniform representation and it is a simulation of the spatial probability distribution of the electrons in a high potential protein molecule. In Fig. 5(b) we show the graph of the RMSE as we increase the resolution of reconstruction. The  $RMSE \leq 1$  condition is not fulfilled even at the resolution  $64 \times 64 \times 64$  which is the original data resolution. In Fig. 5(a) we analyze the graphical plots of the topology-controller for this dataset. We observe that although  $\mathcal{TC}_0 \leq 0.2$  (blue line), our topology criteria,  $\mathcal{TC} \leq 0.2$  (red line), is not fulfilled. This is a typical case when *a priori* knowledge of data helps in fine tuning the topology-controller. The most important features of the Neghip dataset are the connected components. Homological analysis tells us that the number of tunnels and voids (holes) is very low (mostly zero). Changes of the resolution of reconstruction results in the creation or destruction of such structures. This shows up in the graphs in a very emphasized oscillations

for  $\mathcal{TC}_1$  and  $\mathcal{TC}_2$ , due to the usage of the relative error estimator in our topology-controller<sup>4</sup>. In turn this affects also the graph of  $\mathcal{TC}$ . Hence, in such case the usage of the weights ( $\alpha_0 = 1$ ,  $\alpha_1 = \alpha_2 = 0$ ) is suggested. In Fig. 5(c) and (d) we show two renderings, one from the reconstruction with a low resolution Fig. 5(c), and the other with the reconstruction suggested by the topology-controller Fig. 5(d). The clear differences support our statements for the usefulness of the topology-controller in the assessment of reconstruction quality.

## 5. Conclusion and future work

Most of the existing approaches that connect topology with visualization are applied to data given as simplicial complexes. While there are very good reasons to adapt such an approach for rendering, we advocated that an intermediate transformation onto a regular data structure opens up the possibilities for much more sophisticated data processing in general. As such, the need of an optimal resolution of reconstruction becomes obvious.

<sup>4</sup> Similar to the division by a number close to zero

Our work helps towards a better understanding of the reconstruction result without the need of visual inspection. This holds especially for datasets where connected components, tunnels and voids are present. We showed several scenarios from different datasets where a topology-based statistics is very helpful and complementary to reconstruction quality assessment. The topology-controller does not solve definitively the non-trivial process of optimal resolution selection for reconstructing non-uniform point sets. It would be very hard to generalize such concept for every kind of data. On the other hand, our proposed method offers an intuitive approach to handle such problem. It provides the user with intrinsic information about changes in data, hence helping the acceleration of the reconstruction/visualization pipeline.

Although the variational reconstruction is used in this algorithm, the reconstruction module can be replaced by any other reconstruction module. Such versatility makes our topology-controller an important complimentary tool for the improvement of the reconstruction pipeline.

One of the major drawbacks of our algorithm is that we have to explicitly reconstruct the data and compute the Betti numbers. Hence considerable resources are used for such offline preprocessing. Although the focus of this paper was not computational performance, we plan to tackle such issue in our future work. Further future research includes applying the same statistical concepts for time-varying data and investigating the possibility of incorporating the topology controller in the variational reconstruction process.

## Acknowledgment

This work was supported by the Austrian Science Fund (FWF) Grant No. P20134-N13 and the Austrian COMET program.

## Appendix A

We define the 3D Laplacian kernel as follows:

$$\text{Laplacian}(V) = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \quad (5)$$

where  $V$  represents the volume given as a 3D regular grid.

## Appendix B

The RMSE is defined as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_i^M (F(x_i, y_i, z_i) - f_i)^2}{M}} \times \frac{100}{\text{Max}_{value}} \quad (6)$$

where  $F$  is the approximating function,  $f$  are the given values,  $\text{Max}_{value}$  is the maximum value in the given point set and  $M$  is the number of points. For the error estimation in the Laplacian datasets we will compute the error in all regular points (including the points not retained in the Laplacian dataset, i.e.,  $M = N_x \times N_y \times N_z$ ).

## References

Aldroubi, A., Gröchenig, K., 2000. Beurling–Landau-type theorems for non-uniform sampling in shift invariant spline spaces. *J. Fourier Anal. Appl.* 6, 93–103.

- Arigovindan, M., Sühling, M., Hunziker, P.R., Unser, M., 2005. Variational image reconstruction from arbitrarily spaced samples: A fast multiresolution spline solution. In: *Proc. IEEE Trans. Image Process.*, vol. 14, pp. 450–460.
- Bajaj, C.L., Pascucci, V., Schikore, D., 1997. The contour spectrum. In: *Proc. IEEE Visualizat.*, pp. 167–174.
- Bendich, P., Edelsbrunner, H., Kerber, M., 2010. Computing robustness and persistence for images. *IEEE Trans. Vis. Comput. Graph.* 16 (6), 1251–1260.
- Bremer, P.-T., Weber, G.H., Pascucci, V., Day, M., Bell, J.B., 2010. Analyzing and tracking burning structures in lean premixed hydrogen flames. *IEEE Trans. Vis. Comput. Graph.* 16 (2), 248–260.
- Carr, H., Snoeyink, J., van de Panne, M., 2010. Flexible isosurfaces: Simplifying and displaying scalar topology using the contour tree. *Comput. Geometry* 43 (1), 42–58.
- Cerri, A., Biasotti, S., Giorgi, D., 2007. k-Dimensional size functions for shape description and comparison. In: *Int. Conf. Image Anal. Process.*, pp. 795–800.
- Chazal, F., Cohen-Steiner, D., Guibas, L.J., Mémoli, F., Oudot, S., 2009. Gromov-hausdorff stable signatures for shapes using persistence. *Comput. Graph. Forum.* 28 (5), 1393–1403.
- Duchon, J., 1979. Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In: Schempp, W., Zeller, K. (Eds.), *Multivariate Approx. Theory.* Birkhäuser-Verlag, pp. 85–100.
- Edelsbrunner, H., Letscher, D., Zomorodian, A., 2002. Topological persistence and simplification. *Discrete Comput. Geometry* 28 (4), 511–533.
- Feichtinger, H.G., Gröchenig, K., Strohmer, T., 1995. Efficient numerical methods in non-uniform sampling theory. *Numerische Mathematik* 69, 423–440.
- Grishin, D., Strohmer, T., 2004. Fast multi-dimensional scattered data approximation with Neumann boundary conditions. *Linear Algebr. Appl.* 391, 99–123.
- Gyulassy, A., Natarajan, V., Pascucci, V., Bremer, P.-T., Hamann, B., 2005. Topology-based simplification for feature extraction from 3D scalar fields. In: *Proc. IEEE Visualizat.*, pp. 535–542.
- Jang, Y., Botchen, R.P., Lauser, A., Ebert, D.S., Gaither, K.P., Ertl, T., 2006. Enhancing the interactive visualization of procedurally encoded multifield data with ellipsoidal basis functions. *Comput. Graph. Forum.* 25 (3), 587–596.
- Kaczynski, T., Mischaikow, K., Mrozek, M., 2004. *Computational Homology.* Applied Mathematical Sciences, vol. 157. Springer-Verlag.
- Kraus, M., Ertl, T., 2001. Topology-guided downsampling. In: *Proc. Vol. Graph.*, pp. 223–234.
- Matsumoto, Y., 2001. *An Introduction to Morse Theory.* American Mathematical Society.
- Meine, H., Köthe, U., Stelldinger, P., 2009. A topological sampling theorem for robust boundary reconstruction and image segmentation. *Discrete Appl. Math.* 157 (3), 524–541.
- Milnor, J., 1963. *Morse Theory.* Princeton University Press.
- Munkres, J.R., 1984. *Elements of Algebraic Topology.* Addison-Wesley, Redwood City, California.
- Natarajan, V., Pascucci, V., 2005. Volumetric data analysis using Morse–Smale complexes. In: *Proc. Shape Model. Int.*, pp. 322–327.
- Nielson, G.M., 1993. Scattered data modeling. *IEEE Comput. Graph. Appl.* 13, 60–70.
- Ohtake, Y., Belyaev, A.G., Seidel, H.-P., 2004. 3D scattered data approximation with adaptive compactly supported radial basis functions. In: *Proc. Int. Conf. Shape Model. Appl.*, pp. 31–39.
- Schlatter, P., 2001. Direct numerical simulation of laminar-turbulent transition in boundary layer subject to free-stream turbulence. Master's thesis, Royal Institute of Technology, Stockholm.
- Silver, D., Wang, X., 1998. Tracking scalar features in unstructured datasets. In: *Proc. IEEE Visualizat.*, pp. 79–86.
- Stelldinger, P., Latecki, L.J., Siqueira, M., 2007. Topological equivalence between a 3d object and the reconstruction of its digital image. *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (1), 126–140.
- Takahashi, S., Nielson, G.M., Takeshima, Y., Fujishiro, I., 2004. Topological volume skeletonization using adaptive tetrahedralization. In: *Proc. Geometric Modell. Process.*, pp. 227–236.
- Thévenaz, P., Blu, T., Unser, M., 2000. Interpolation revisited. *IEEE Trans. Med. Imaging* 19 (7), 739–758.
- Vućini, E., Möller, T., Gröller, M.E., 2008. Efficient reconstruction from non-uniform point sets. *The Visual Computer*, vol. 24. Springer, Berlin/Heidelberg (7–9) 555–563.
- Vućini, E., Möller, T., Gröller, M.E., 2009. On visualization and reconstruction from non-uniform point sets using b-splines. In: *Proceedings of Eurographics/ IEEE-VGTC Symposium on Visualization*, vol. 28, pp. 1007–1014.
- Weber, G.H., Dillard, S.E., Carr, H., Pascucci, V., Hamann, B., 2007. Topology-controlled volume rendering. *IEEE Trans. Vis. Comput. Graph.* 13 (2), 330–341.
- Weber, G.H., Bremer, P.-T., Day, M.S., Bell, J.B., Pascucci, V., 2009. Feature tracking using reeb graphs. In: *Proc. Topol.-based Methods Visualizat.*