

Enhancing the Reconstruction from Non-uniform Point Sets using Persistence Information

Erald Vućini

VRVis Center for Virtual Reality and Visualization Research
vucini@vrvis.at

Abstract. In this paper we propose an efficient method for selecting the reconstruction resolution of non-uniform representations. We analyze the topological difference between reconstructions based on Topological Persistence information and define a distance for quantifying such information. We compute the Persistence information with a state-of-the-art method and report in detail the characteristics of the proposed algorithm. We evaluate our method in different scenarios and compare to previous contributions. Our proposed method offers faster and more reliable results in an effort to improve the reconstruction process and to reduce the necessity for visual inspection.

Keywords: Reconstruction; Topology; Persistence; Bottleneck Distance

1 Introduction

Nowadays, the number of applications that provide non-uniform (irregular) data is increasing steadily. Examples vary from astronomical, Doppler and ultrasound measurements, to particle or numerical simulations in physical sciences. Non-uniform data representations offer a way of adapting the measure location according to the importance of the data. However, most of the techniques that deal with the analysis and processing of data are fitted to uniform (regular) data. The transform of non-uniform representations to uniform ones, is a viable option, for the better understanding and analysis of such data. In this paper we focus on the process of reconstruction from non-uniform to uniform representations.

The main problem in the reconstruction from non-uniform representations is the selection of the proper resolution of reconstruction. This will be the central question we will try to answer in this paper. The problem of selection of reconstruction resolution can be translated into a trade-off finding problem between accuracy and memory efficiency. A coarse (low) resolution of reconstruction requires less memory consumption, but will smooth the signal in areas with sharp transitions, resulting in visual artifacts or high errors. On the other side, a too fine (high) resolution of reconstruction will introduce memory overheads while achieving a better accuracy.

Another important issue, in the reconstruction process is the evaluation of the quality of reconstruction. 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 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. Vućini and Kropatsch [17] proposed to reduce the necessity for visual inspection by using topological information derived from Homology analysis. A schematic view of this reconstruction pipeline is displayed in Fig. 1.

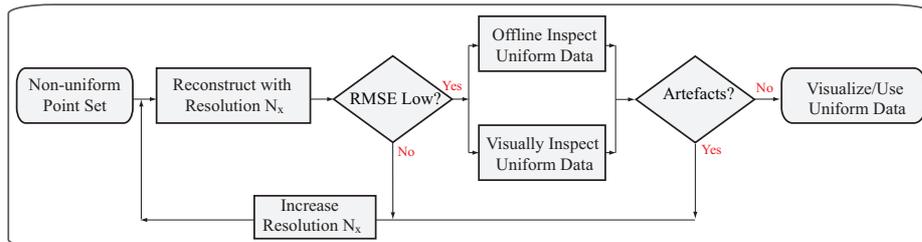


Fig. 1. Schematic view of a pipeline for the reconstruction of non-uniform point sets to uniform representations when the target resolution is unknown. The offline inspection step represents the option to decide over the quality of reconstructions using additional information and without the need of visual inspection.

In this paper, we build upon the work of Vućini and Kropatsch [17]. We also derive a topological signature from the analysis of data. Instead of homology we use persistent homology (or simple Persistence) in our analysis workflow. The usage of Persistence allows us to compute a more robust topological signature, in a more efficient way due to recent developments in persistence computation algorithms [20]. The topological information together with error measurements improves the quality assessment of the reconstruction and reduces the need of visual inspection.

In Section 2 we give a short summary of related work. We introduce our topology controller in Section 3, and explain the main modules of the proposed algorithm. In Section 4 we show results w.r.t. the usage of the topological signature and assess the value of its applicability. Finally, conclusions are drawn and some ideas are layed out and discussed as future work in Section 5.

2 Related Work

Following we give an overview of work related to the proposed approach, namely, with regard to reconstruction and topology analysis of data.

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 ([11]). Perhaps the most popular approach for approximating non-uniform data is based on Radial Basis Functions (RBFs). They have been used in surface ([14]) as well as volumetric ([13]) approximation and reconstruction techniques. Arigovindan et al. [2] proposed to use B-splines in a multi-grid framework for the reconstruction of non-uniform 2D data. Vućini et al. [18] extended these ideas to 3D volumes and large datasets.

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 ([1]). 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 ([15]). Vuçini et al. [18] proposed the usage of the σ_{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.

Due to the complexity of the data, techniques for providing a simplified view are required in fields such Visualization and Graphics. Topology analysis has been successfully linked to fields related to isosurface selection [3], topological downsampling and simplification ([12]), topology-guided analysis and navigation in scalar and time varying data ([5]), and feature tracking and evolution ([21]). Carr et al. [6] have presented a generalized framework consolidating the theory and application of the contour spectrum concept. Most of the above-mentioned works have concentrated in reporting topological information related to 0-dimensional homology, i.e., connected components. Topological persistence information has been also used for shape comparison and feature classification ([7], [8], [4]). In our previous work ([17]), homology information was used to derive a topological signature. The main drawback of the method, is the necessity to compute the homology information for each superlevel set in order to derive important information. In the current approach the usage of Persistence overpasses this drawback, resulting in lower computation times and in the same time providing a more compact representation.

All the above-mentioned methods provide extensive information, which is difficult to interpret without the appropriate statistical analysis. Similar to [17], in the proposed method we use statistical topological information for guiding the selection of resolution of reconstruction. As a result, our framework gives important cues that reduce the necessity of human's visual inspection of the data.

3 Topology-based Analysis

Our proposed algorithm consists of two main modules: 1) the variational reconstruction module, and 2) the module that derives the statistical Persistence-based 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 Persistence-based topology-controller (\mathcal{PC}).

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{S} of sample points, $\mathbf{s}_i = (x_i, y_i, z_i)$, $i = 1, 2, \dots, M$, the B-spline approximation is formulated in a way that it approximates the points set in a resolution (N_x, N_y, N_z) of the axis-aligned bounding box. 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 ([16]).

The key idea of the variational reconstruction is to build a linear system and to solve it by minimizing a derived cost function. Once the linear system is solved, we can estimate the approximating function $f(\mathbf{s})$ (a C^2 -continuous function) at any position $\mathbf{s} \in V$, 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 [18].

3.2 Persistence-based Topology Analysis

We shortly introduce Persistence, focusing on \mathbb{Z}_2 homology ([10]). Given a topological space \mathbb{X} and a *filtering function* $f : \mathbb{X} \rightarrow \mathbb{R}$, Persistence performs a topological exploration along a filtration, i.e., a nested sequence of subsets $X_1 \subseteq X_2 \subseteq \dots \subseteq X_n = X$, usually induced by considering the sublevel sets of the filtering function. The algorithm captures the birth and death times of homology classes of the sublevel set as it grows along the filtration. By birth, we mean that a homology class comes into being; by death, we mean it either becomes trivial or becomes identical to some other class born earlier. The persistence, or lifetime of a class, is the difference between the death and birth times. Homology classes with larger persistence reveal information about the global structure of the space \mathbb{X} , as described by the function f .

3.3 Persistence-based Algorithm (PbA)

Similar to [17], the proposed algorithm takes as input a non-uniform point set (\mathbf{S}), a minimum (N_{min}) and maximum (N_{max}) resolution of reconstruction, and a resolution step Δ (see Algorithm 1). In difference from [17] we do not require the number of superlevel sets needed to reconstruct. The number of superlevel sets was previously required by the Homology module, due to the fact that homology computation runs only on binary data, e.g., object and background. Persistence, offers us the possibility to create a filtration along the values of the data.

Algorithm 1 PbA(\mathbf{S} , N_{min} , N_{max} , Δ)

- 1: **for** $i = N_{min}$ to N_{max} **do**
 - 2: determine V_i by solving variational reconstruction for \mathbf{S} on resolution i
 - 3: compute Persistence $\mathcal{P}_{V_i}^0$, $\mathcal{P}_{V_i}^1$ and $\mathcal{P}_{V_i}^2$
 - 4: build Persistence frequency histograms $\mathbf{PH}_{V_i}^0$, $\mathbf{PH}_{V_i}^1$ and $\mathbf{PH}_{V_i}^2$
 - 5: $i = i + \Delta$
 - 6: **end for**
 - 7: compute the topology controller (\mathcal{PC})
-

The iterative algorithm starts with determining the volume V_i , from the approximating function estimated from the variational reconstruction of the non-uniform point set \mathbf{S} (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 Δ). 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. In (line 3) we compute the persistence

of the reconstructed volumes. As a result, we obtain a set of pairs for each dimension, i.e., $\mathcal{P}_{V_i}^0$, $\mathcal{P}_{V_i}^1$ and $\mathcal{P}_{V_i}^2$. Each pair represents a persistent topological feature and is given as two functional values, i.e., (birth, death). Following, we can build the Persistence frequency histograms $\mathbf{PH}_{V_i}^0$, $\mathbf{PH}_{V_i}^1$ and $\mathbf{PH}_{V_i}^2$ (line 4). The frequency histograms span the data range, e.g., in many data examples [0,4095], and measure the number of persistent topological features that are alive at a specific functional value. $\mathbf{PH}_{V_i}^\tau$ can be considered as the τ -dimensional persistence signature of the point set \mathbf{S} in resolution i .

After computing the Persistence frequency histograms for each resolution, we can define the topology controller $\mathcal{P}\mathcal{C}(i)$ as:

$$\mathcal{P}\mathcal{C}(i) = \frac{1}{\alpha_0 + \alpha_1 + \alpha_2} \sum_{\tau=0}^2 \alpha_\tau \cdot \frac{\|\mathbf{PH}_{V_i}^\tau - \mathbf{PH}_{V_{N_{max}}}^\tau\|}{\|\mathbf{PH}_{V_{N_{max}}}^\tau\|} \quad (1)$$

where the weights (coefficients) $\alpha \in \{0, 1\}$ control the impact of the respective τ -dimensional persistence statistics (\mathbf{PH}^τ) on the topology controller. In simpler words, $\mathcal{P}\mathcal{C}(i)$ computes the relative error of $\mathbf{PH}_{V_i}^\tau$ with regard to $\mathbf{PH}_{V_{N_{max}}}^\tau$, which is the persistence signature of the point set \mathbf{S} in the maximum resolution.

4 Implementation and Results

Our test platform is an Intel i7 CPU @ 2.67GHz with 12GB of RAM. All the algorithms are developed as single threaded hence only one processor core is used. Analog to [17], we tested our framework with 3D data sets based on non-uniform point sets as well as Cartesian grids. For detailed information on the datasets refer to [19].

In our framework, we analyze graphical plots of $\mathcal{P}\mathcal{C}$ with regard to a changing resolution and we set default weight values ($\alpha_0 = \alpha_1 = \alpha_2 = 1$) in Eq. 1. We attach to these plots also the graphs of $\mathcal{P}\mathcal{C}_\tau$, which measure the τ -dimensional homological statistics. $\mathcal{P}\mathcal{C}_\tau$ is derived from $\mathcal{P}\mathcal{C}$ by setting the respective α_τ equal to one and the other two weights equal to zero. Vućini and Kropatsch [17], suggested a threshold of 0.2 for the topology controller, in order to find the optimal resolution of reconstruction. While this is a viable approach, in this work we followed a different strategy. We let the user define upon the plots of the topology controller, which the resolution that better suits her. In this way, we avoid hard thresholding, which is not optimal when dealing with different data types.

In Fig. 2 we show renderings of the point set and reconstructions from the Cooling Jacket dataset. The dataset was generated at AVL List GmbH in order to evaluate a cooling jacket design for a four cylinder diesel engine. This stationary flow simulation incorporates a heat transport solution in order to predict critical temperature regions within the engine. The original dataset is specified on an unstructured grid. The non-uniform point set consists of 1,537,898 points encoding the pressure of the flow data.

The topology of the data is very crucial for such an industrial simulation, and our algorithm should be able to distinguish between the cases displayed in Fig. 2 b-d). In Fig. 3(a), we display the plots of the topology controller derived from our proposed algorithm. We see that the (red) line representing $\mathcal{P}\mathcal{C}$, stabilizes after the resolution $N_x = 410$. Hence, this is the value that we select as an optimal reconstruction resolution for the Cooling Jacket dataset. Visually comparing the renderings in Fig. 2 and 3(b), we can also confirm that the reconstruction done with the selected resolution is topologically more similar to the input point set, than the ones done with lower resolutions.

In Table 1 we provide results of our framework for different datasets and also compare with the methods proposed in [18] and [17]. While our proposed resolutions are different from the ones suggested in [18], they are similar to the ones proposed in [17].

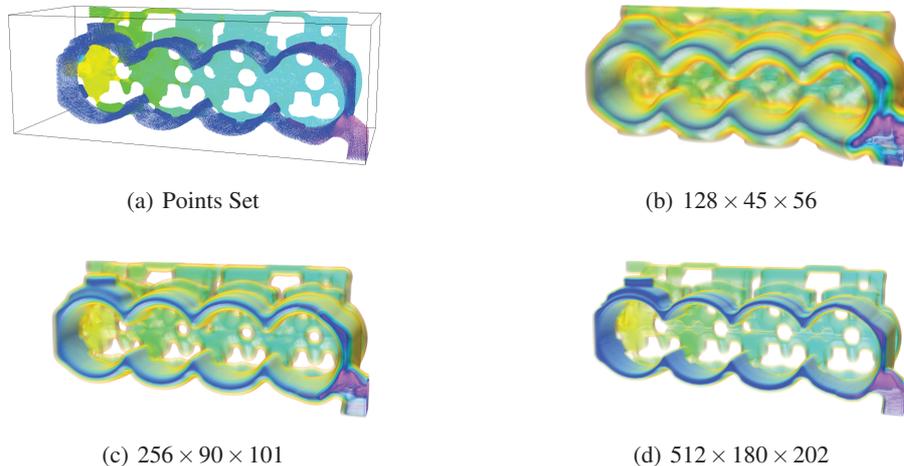


Fig. 2. Cooling Jacket (Pressure) dataset: a) dataset consisting of 1,537,898 non-uniform points, b-d) reconstructions with resolution $128 \times 45 \times 51$, $256 \times 90 \times 101$ and $512 \times 180 \times 202$ respectively.



Fig. 3. Cooling Jacket (Pressure) dataset: a) Persistence-based controller, b) reconstruction with resolution $410 \times 144 \times 162$. The hair-line in the Fig. a) highlights the values of the Persistence Controllers for the selected resolution ($N_x = 410$).

For computing the Persistence information we use the algorithm developed by Wagner et al. [20]. For obtaining the reconstruction we use the algorithm developed by Vuçini et al. [18]. The complexity of Persistence algorithm is cubic, while the variational reconstruction algorithm has a linear complexity. However, in our framework we observed a linear complexity from both the modules. Computation times of the proposed framework, and comparison to [17] are given in Table 2. Note that, with the usage of the Persistence module we significantly decrease the computation times.

5 Conclusions and Future Work

In this work we presented an efficient method for selecting the resolution of reconstruction for non-uniform point sets. Building this work on our previous contribution, we

provide a more efficient method based on the concept of Persistence. We demonstrated our results in comparison with previous contributions.

One of the reasons of using persistence, was to be able to use the bottleneck distance [9], as a measure for quantifying the difference between reconstructions in various resolutions. Along this work, we tested the proposed algorithm with both the bottleneck and Hausdorff distances. However, both preliminary testings resulted in a non-converging distance. In our future work, we want to analyze in more detail, the reasons underlying such fact. We believe, this can be related with the way our distance computing algorithm handles essential and non-essential classes in the persistence diagram. Another reason could be related to the possible normalizations, that have to be done in order to be able to compare different persistence diagrams.

In our future work, we also want to enable the usage of persistence-based topological signatures in visualization and graphics, e.g., in designing transfer functions or highlighting topological features of interest.

Table 1. Comparisons of reconstruction resolutions for different non-uniform datasets as proposed by [18] and our proposed Topology controller.

Dataset		Previous [18]		Previous [17]		Proposed	
Name	Points	Resolution	RMSE	Resolution	RMSE	Resolution	RMSE
Oil	29,094	38x40x38	0.19	41x44x41	0.16	43x45x43	0.15
Natural Convection	68,921	61x61x61	0.63	55x55x55	0.71	58x58x58	0.68
Synthetic Chirp	75,000	64x64x64	1.12	58x58x58	1.18	60x60x60	1.16
Bypass	7,929,856	766x92x192	0.61	880x106x220	0.49	860x103x216	0.55
Blunt-Fin	40960	93x36x25	1.14	102x39x28	1.11	98x38x26	1.12
Cooling Jacket	1,537,898	256x90x101	0.92	400x140x157	0.86	410x144x162	0.83

Table 2. Comparison of times for the computation of the controllers for different datasets and settings, as reported by Vućini and Kropatsch[17] and the proposed method. (*) Timings for the results in the 'Previous' column are estimated on a PC, running on an Intel Dual Core 2.70 GHz processor machine with 6GB of RAM (similar single-core performance).

Dataset					Times (min)	
Name	Points	N_{min}	N_{max}	Δ	Previous [17](*)	Proposed
Neghip	52,428	16	80	1	60.45	1.47
Natural Convection	68,921	16	64	1	107.13	0.63
Aneurism	419,430	32	160	4	40.77	8.12
Cooling Jacket(Pressure)	1,537,898	32	512	2	7215.32	282.13
Bypass	7,929,856	256	1024	16	6712.80	257.98

6 Acknowledgement

This work was supported by the Austrian Science Fund (FWF) grant no. P20134-N13 and the Austrian COMET program. The author thanks Dr. Andrea Cerri and Prof. Walter Kropatsch for the useful discussions related to the topic of this paper.

References

1. Akram Aldroubi and Karlheinz Gröchenig, *Beurling-Landau-type theorems for non-uniform sampling in shift invariant spline spaces*, Journal of Fourier Analysis and Applications **6** (2000), 93–103.

2. Muthuvel Arigovindan, Michael Sühling, Patrick R. Hunziker, and Michael Unser, *Variational image reconstruction from arbitrarily spaced samples: A fast multiresolution spline solution*, Proceedings of IEEE Transactions on Image Processing, vol. 14, 2005, pp. 450–460.
3. Chandrajit L. Bajaj, Valerio Pascucci, and Daniel Schikore, *The contour spectrum*, Proceedings of IEEE Visualization, 1997, pp. 167–174.
4. Paul Bendich, Herbert Edelsbrunner, and Michael Kerber, *Computing robustness and persistence for images*, Proceedings of IEEE Visualization, 2010, pp. 1251–1260.
5. Peer-Timo Bremer, Gunther H. Weber, Valerio Pascucci, Marc Day, and John B. Bell, *Analyzing and tracking burning structures in lean premixed hydrogen flames*, IEEE Trans. Vis. Comput. Graph. **16** (2010), no. 2, 248–260.
6. Hamish Carr, Jack Snoeyink, and Michiel van de Panne, *Flexible isosurfaces: Simplifying and displaying scalar topology using the contour tree*, Computational Geometry **43** (2010), no. 1, 42–58.
7. Andrea Cerri, Silvia Biasotti, and Daniela Giorgi, *k-dimensional size functions for shape description and comparison*, International Conference on Image Analysis and Processing, 2007, pp. 795–800.
8. Frédéric Chazal, David Cohen-Steiner, Leonidas J. Guibas, Facundo Mémoli, and Steve Oudot, *Gromov-hausdorff stable signatures for shapes using persistence*, Comput. Graph. Forum **28** (2009), no. 5, 1393–1403.
9. David Cohen-Steiner, Herbert Edelsbrunner, and John Harer, *Stability of persistence diagrams*, Discrete and Computational Geometry **37** (2007), no. 1, 103–120.
10. Herbert Edelsbrunner and John Harer, *Computational topology, an introduction.*, American Mathematical Society, 2010.
11. Hans G. Feichtinger, Karlheinz Gröchenig, and Thomas Strohmer, *Efficient numerical methods in non-uniform sampling theory*, Numerische Mathematik **69** (1995), 423–440.
12. Attila Gyulassy, Vijay Natarajan, Valerio Pascucci, Peer-Timo Bremer, and Bernd Hamann, *Topology-based simplification for feature extraction from 3D scalar fields*, Proceedings of IEEE Visualization, 2005, pp. 535–542.
13. Yun Jang, Ralf P. Botchen, Andreas Lauser, David S. Ebert, Kelly P. Gaither, and Thomas Ertl, *Enhancing the interactive visualization of procedurally encoded multifield data with ellipsoidal basis functions*, Computer and Graphics Forum **25** (2006), no. 3, 587–596.
14. Yutaka Ohtake, Alexander G. Belyaev, and Hans-Peter Seidel, *3D scattered data approximation with adaptive compactly supported radial basis functions*, Proceedings of International Conference on Shape Modeling and Applications, 2004, pp. 31–39.
15. Peer Stellinginger, Longin Jan Latecki, and Marcelo Siqueira, *Topological equivalence between a 3d object and the reconstruction of its digital image*, IEEE Trans. Pattern Anal. Mach. Intell. **29** (2007), no. 1, 126–140.
16. Philippe Thévenaz, Thierry Blu, and Michael Unser, *Interpolation revisited*, IEEE Transactions on Medical Imaging **19** (2000), no. 7, 739–758.
17. Erald Vuçini and Walter G. Kropatsch, *On the search of optimal reconstruction resolution*, Pattern Recognition Letters (2011), In Press.
18. Erald Vuçini, Torsten Möller, and M. Eduard Gröller, *On visualization and reconstruction from non-uniform point sets using b-splines*, Proceedings of Eurographics/ IEEE-VGTC Symposium on Visualization, vol. 28, 2009, pp. 1007–1014.
19. Erald Vucini, *On visualization and reconstruction from non-uniform point sets*, Ph.D. thesis, Vienna University of Technology, 2009.
20. Hubert Wagner, Chao Chen, and Erald Vuçini, *Efficient computation of persistent homology for cubical data*, Proceedings of Workshop on Topology-based Methods in Data Analysis and Visualization, 2011, To appear in Mathematics + Visualization series by Springer.
21. Gunther H. Weber, Peer-Timo Bremer, Marcus S. Day, John B. Bell, and Valerio Pascucci, *Feature tracking using reeb graphs*, Proceedings of Topology-based Methods in Visualization, 2009.