Topology Preserving Compressor

Nathan Gorski, Xin Liang, Hanqi Guo, Lin Yan, and Bei Wang

Overview

• Lossy compression reduces the size of a file while allowing a controlled amount of error.

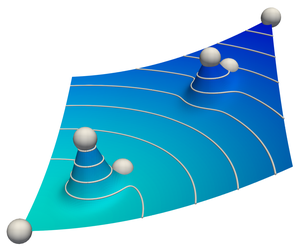
• Lossy compressors distort geometric and topological information in data, such as the contour tree.

• Our framework modifies any lossy compressor to preserve the contour tree and maintain a specified error bound.

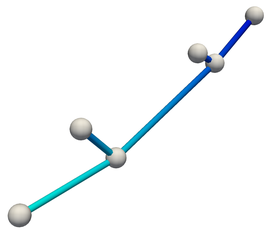
Contour Tree

• A contour is a connected set of points with the same function value (same height, temperature, etc.)

• The contour is a summary of data formed by collapsing contours into a single point. It has applications in several fields, including astronomy, graphics, and medicine.



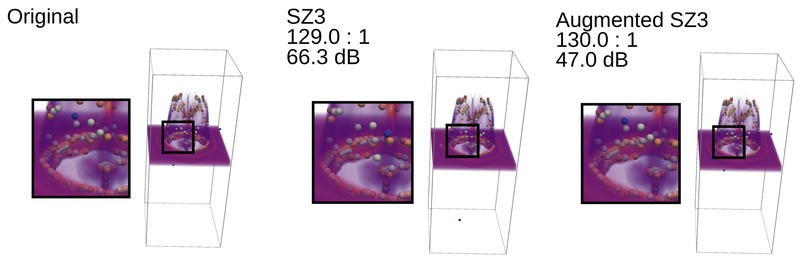
contours



contour tree

Problem

• Existing lossy compressor distort the contour tree, even with a very small error bound.

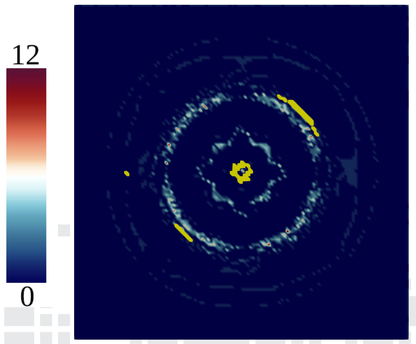


Vertices of the contour tree of an ionization front before compression (left), after compression with the SZ3 compressor (middle), and after compression with SZ3 augmented with our framework (right). The middle figure uses an error bound of 3.75e-4. The right figure uses an error bound of 0.012. The data range is [0,1]. While the compressed file sizes are equal, our framework perfectly preserves the vertices while SZ3 does not, despite its very low error bound.

Our Strategy

• Iteraively calculate error bounds for each individual point that, if maintained, preserve the contour tree and global error bound, similar to Yan et al. [2]

• Calculate discrete adjustments that must be made to the output of any compressor to maintain these error bounds. Store them using a novel variable-precision encoding scheme. More precision is used only for points that need it.



Log of precision used to store points for a slice of the ionization front dataset.

Yellow points required too much precision and are stored losslessly.

Results

Guarantees

• The contour tree is always preserved after edges (a,b) are removed where |f(a)-f(b)| is less than a given parameter ε.

• The specified error bound is always maintained.

Comparison to state-of-the-art

• We compare to topology preserving compressors TopoQZ [1] and TopoSZ [2].

• When modifying the SZ3 compressor, our framework produces smaller file sizes than TopoQZ and TopoSZ with similar reconstruction quality.

Dataset	File Size	Ours	TopoSZ	TopoQZ
Earthquake	28.2M	**277K**	562.9K	398.9K
Ionization	40.6M	**312.3K**	1.6M	534.6K
Isabel	105M	**1.3M**	2.7M	3.5M
Miranda	302M	**1.5M**	3.1M	4.2M
QMCPack	4.4M	**70.5K**	157.6K	128.1K
Tangaroa	27M	**913.2K**	1.1M	990.1K

Compressed file sizes for augmented SZ3 (ours) vs TopoSZ and TopoQZ.
ε = 4% of range. Error = 1.2% of range.

References

[1] Soler, Maxime, Mélanie Plainchault, Bruno Conche, and Julien Tierny. "Topologically controlled lossy compression." In 2018 IEEE Pacific Visualization Symposium (PacificVis), pp. 46-55. IEEE, 2018.

[2] Yan, Lin, Xin Liang, Hanqi Guo, and Bei Wang. "TopoSZ: Preserving Topology in Error-Bounded Lossy Compression." IEEE Transactions on Visualization and Computer Graphics (2023).

This research was partially supported by NSF OAC-2313124.

THE UNIVERSITY OF UTAH

SCI www.sci.utah.edu