Analysis of Large-Scale Scalar Data Using Hixels

Topology for Statistical Modeling of Petascale Data

David Thompson, Janine Bennett, Philippe Pébay
Sandia National Laboratories, P.O. Box 969, MS 9159, Livermore, CA 94551

Joshua Levine, Peer-Timo Bremer, Attila Gyulassy, Valerio Pascucci
Scientific Computing & Imaging Institute, University of Utah, Salt Lake City, Utah 84112

Abstract

One of the greatest challenges for today’s visualization and analysis communities is the massive amounts of data generated from state of the art simulations. Traditionally, the increase in spatial resolution has driven most of the data explosion, but more recently ensembles of simulations with multiple results per data point and stochastic simulations storing individual probability distributions are increasingly common. This work introduces a new data representation for scalar data, called hixels, that stores a histogram of values for each sample point of a domain. The histograms may be created by spatial down-sampling, binning ensemble values, or polling values from a given distribution. In this manner, hixels form a compact yet information rich approximation of large scale data. In essence, hixels trade off data size and complexity for scalar-value “uncertainty”.

Based on this new representation we propose new feature detection algorithms using a combination of topological and statistical methods. In particular, we show how to approximate topological structures from hixel data, extract structures from multi-modal distributions, and render uncertain isosurfaces. In all three cases we demonstrate how using hixels compares to traditional techniques and provide new capabilities to recover prominent features that would otherwise be either infeasible to compute or ambiguous to infer. We use a collection of computer tomography data and large scale combustion simulations to illustrate our techniques.