Re-projection of Terabyte-Sized Images

Peter Bajcsy1, Antoine Vandecreme1 and Mary C. Brady1

1Software and Systems Division
Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899

Abstract: This work addresses the problem of re-projecting a terabyte-sized 3D data set represented as a set of 2D Deep Zoom pyramids. In general, a re-projection for small 3D data sets is executed directly in RAM. However, RAM becomes a limiting factor for terabyte-sized 3D volumes formed by a stack of hundreds of megapixel to gigapixel 2D frames. We have analyzed and benchmarked five methods to perform the re-projection computation in order to overcome the RAM limitation.

How Does One Inspect Terabyte-sized 3D Images From Multiple Viewpoints?

We address the problem of enabling interactive visualization of terabyte-sized 3D images from multiple viewpoints in a web browser.

Motivation: With the current limitations of desktop computers in terms of RAM, storage and computation power, the Deep Zoom paradigm [1] has been frequently adopted for visualization of large 2D image data sets. In this work, we focus on re-projecting 3D data sets to deliver multiple 2D views using the Deep Zoom paradigm.

3D Re-projection Algorithms

Single Computer: We compared three implementations to generate pyramids of re-projection. The first one is a re-projection in RAM where the input 3D volume is in RAM and the output 2D frames are generated one after another. The second one is a re-projection based on accessing disk and RAM holding one 2D frame of the input 3D volume and one of the output 2D volume. The re-projection uses file I/O to read input 2D frames incrementally and to copy one columnrow from input to output. The third one is a re-projection based on copying values from input pyramid tiles to output pyramid tiles where RAM holds only one row of input tiles and one row of re-projected tiles at the highest resolution.

Multiple Computers: We compared two Map/Reduce algorithms applied to the image set to image set and pyramid set to pyramid set representations.

Experimental Hardware

The physical sample was provided by courtesy of Prof. Michael Notis, Lehigh University and Prof. Aaron Shugar, Buffalo State College. The physical sample was imaged by courtesy of Dr. Nicholas Ritchie and Dr. John Henry Scott from NIST.

Summary

Our memory complexity analysis and measured time benchmarks documented that (a) the Hadoop re-projection algorithms enabled handling terabyte sized images on computer cluster/cloud platforms, and (b) the Hadoop algorithm operating on a pyramid set presentation achieves similar utilization of RAM and processors per node as the two algorithms designed for a single computer operating either on pyramid set (PS_TILE FILE) or on frame set (IS_FILE).

Our contributions lie in (a) designing a new approach to 3D re-projection of terabyte-sized images from a set of Deep Zoom pyramids, (b) characterizing computational memory complexities of five re-projection algorithms, and (c) maximizing the speed-up of 3D re-projection computations on Hadoop computer clusters by optimal selection of configuration parameters.


Disclaimer: Any mention of commercial products or reference to commercial organizations is for information only; it does not imply recommendation or endorsement by NIST nor does it imply that the products mentioned are necessarily the best available for the purpose.