Evaluation of Data Layout Effects on Processing High Definition Images

Alex Sherman, Columbia University

Valerio Pascucci, Lawrence Livermore National Laboratory

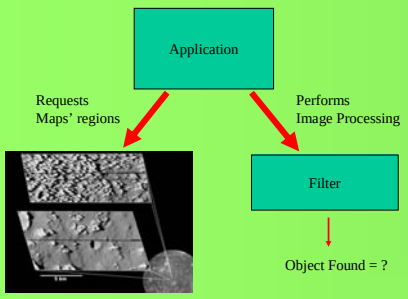
Aug 11, 2005

Introduction

- The speed of processing large high definition images is heavily determined by how quickly the image data can be loaded into memory
- The data layout of an image is crucial in determining the access pattern and therefore the speed of loading the data

Example Application

- Finding a small object on a large high definition satellite map
- The application requests map regions at various resolutions and performs filtering to determine if the region contains the sought object



Proposed Layouts

- Row Major –typical layout. In 2D images, for instance, the image file first contains row 1, then row 2, etc…
- Lebesgue space filling curve (also known as Z-order)
- Hierarchical Lebesgue layout (a.k.a. HZ-order)

Lebesgue Curve: Z-order

- Follows Z-shaped curve to index the data
- Above is a picture of the Z-order curve used to index 2-dimensional data at progressively higher resolutions. At each level of successively finer resolution a vertex of a Z-shape is replaced by a Z-shape ½ the size of the previous level.

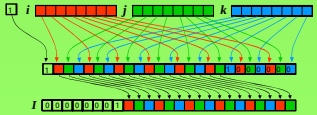
Advantages of Lebesgue Curve

- For a given resolution level easy to convert d-dimensional data into one-dimensional indexing scheme. (Bits used to index each dimension are interleaved to form the one-dimensional index)
- Exploits geometric locality at higher resolutions for sequential and therefore faster disk access

Hierarchical Z order (HZ-order)

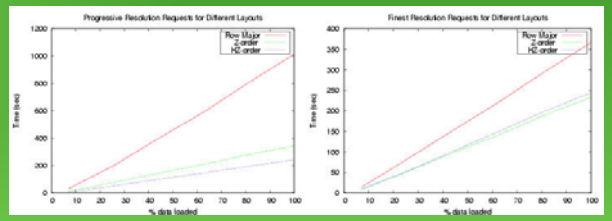
- Proposed by Pascucci and Frank. Implemented by Scorzelli.
- Indexes Z-curve resolution levels in hierarchical order from coarser to finer.
- Maintains the same geometric locality for each Z-curve resolution level
- Beneficial for progressive resolution requests. (E.g. the "object search" application may first attempt to perform filtering on a coarser resolution)

SIMPLE BIT MANIPULATIONS ALLOW TO COMPUTE THE ADDRESS OF A DATA SAMPLE IN HIERARCHICAL Z-ORDER FROM THE ORIGINAL (i,j,k) ROW-MAJOR ARRAY INDEX. THIS ALLOWS TO COMBINE THE COARSE-TO-FINE PROGRESSIVE TRAVERSAL WITH FAST DIRECT DATA ACCESS.

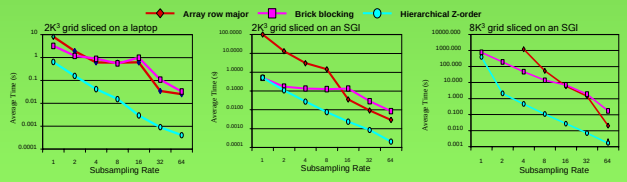


Evaluation

We evaluated the speed of loading a set of random samples from an 8GB 3D image. We compared three data layouts for both progressive resolution and the finest resolution requests



- Both Z and HZ-order significantly outperform the standard Row Major mode representation
- HZ-order also outperforms Z-order for progressive requests but is slightly worse for requests for the finest resolution



Practical performance tests show orders of magnitude improvement with respect to previous storage layouts and scalability with respect to both data size and computing resources available.

This work was performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48