An Efficient Implementation of a Local Binning Algorithm for Digital Elevation Model Generation of LiDAR/ALSM Dataset

Han Suk Kim (UCSD), J Ramon Arrowsmith (ASU), Crosby, Christopher J. (ASU), Efrat Frank (SDSC), Viswanath Nandigam (SDSC), Jeffrey Conner (ASU), Ashraf Memon (SDSC), Scott B. Baden (UCSD), Chaitan Baru (SDSC)

The GEON LiDAR Workflow (GLW) project (http://www.geongrid.org/science/lidar.html) has enabled geoscientific communities to process and download large LiDAR (Light Distance And Ranging), or ALSM (Airborne Laser Swath Mapping) datasets. The high density and random distribution of LiDAR datasets make production of digital elevation models (DEMs) computationally challenging. DEMs are essential representations of the landscape that are imported in various off-the-shelf geoscientific tools.

So far, the GLW has utilized the regularized spline with tension interpolation algorithm for the DEM generation from GRASS GIS successfully for small (<1.6 million points). However, it is compute-intensive; in order to process 5 million LiDAR points, it takes about 50 minutes on the current single processor Linux machines in a GEON cluster.

As a solution to the problem, a local binning algorithm which quickly generates a DEM is employed. The algorithm utilizes the elevation information from only local reference points, the points inside of a circular search area with user specified radius. With the local points, five values are computed for each node in a grid: 1) the minimum, 2) maximum, 3) mean, and 4) inverse distance weighted mean of the local points, and 5) the number of points in the search area). If the number of points in the search radius is 0, the node is assigned a null value. The noble implementation technique can produce a grid containing those five values within O(N) time, where N denotes the size of the point cloud used. In addition, the space cost for this implementation is O(M), where M denotes the size of the grid. This implementation runs over 5 million points in 30 seconds in a single machine, which is about 100 times speedup relative to the $O(N^2)$ spline or related methods.

Along with the original implementation, an out-of-core (memory) version of the local binning algorithm has also been developed. This implementation exploits secondary storage for saving intermediate results when the size of a grid exceeds that of memory. The large grid is split into several pieces so each piece can be fit in main memory. During the computation, the pieces reside in secondary storage and each piece is fetched to main memory one at a time when needed. Although the additional I/O operations give more overhead, the out-of-core version can compute a DEM from 150 million points (in about 20 minutes), which is impossible in the spline interpolation and also in the original local binning algorithm implementation.

In addition to the performance improvements, the local binning algorithm computes useful products. The maximum values track the highest elevation points, which are usually recorded at the top of vegetation areas. On the other hand, the minimum values represent the information bounced from the ground, which describes the ground surfaces more accurately (although this is not as effective as true vegetation filtering, but is useful for meter-scale shrubbery removal for which there will not usually be laser penetration. Lastly, the density map of an area is useful when determining the accuracy of interpolation and assessing the appropriate grid resolution.

This suite of DEM products provides geoscientists rich information complementary to those of the current spline algorithm if they are computable.

In the near future, we will parallelize this algorithm, which will reduce the execution time further. Then, the parallelized and out-of-core implementation will be able to routinely process thousands of requests over broad areas of interest in the GEON LiDAR Workflow.