Intelligent Reduction of Sweep Survey/LIDAR Large Data Sets

By: Mr. Michael Grounds, Beacon Resources
Address: 10152 Long Meadow Rd., Madison, Al. 35758
Phone: (256) 771-0014 Fax: (256) 325-0051
e-mail: **mike@riversrus.com**
Web site: **www.riversrus.com**

Overview

High resolution hydrographic and topographic survey systems generate tremendous volumes of detailed data allowing highly accurate Digital Terrain Models (DTM) to be created. However, engineering analysis tools are frequently limited, although to a lesser degree than in the past, by computer memory and file size constraints. This results in analysis tools which cannot efficiently handle these high-resolution data sets. Frequently the data set size is reduced through the use of arbitrary thinning algorithms such as gridding the data and then reducing the grid resolution, or by simply discarding every third point, fourth point, etc. Ideally, data set size is reduced by intelligent thinning, discarding only those points which do not affect, or only minimally effect, the details of the DTM. In this way points where the detail is needed are retained and points which do not affect the model are discarded. This paper describes the intelligent thinning algorithm, describes the software program workflow, shows algorithm performance for a sample data set, and finally summarizes some of the distinguishing features of this approach.

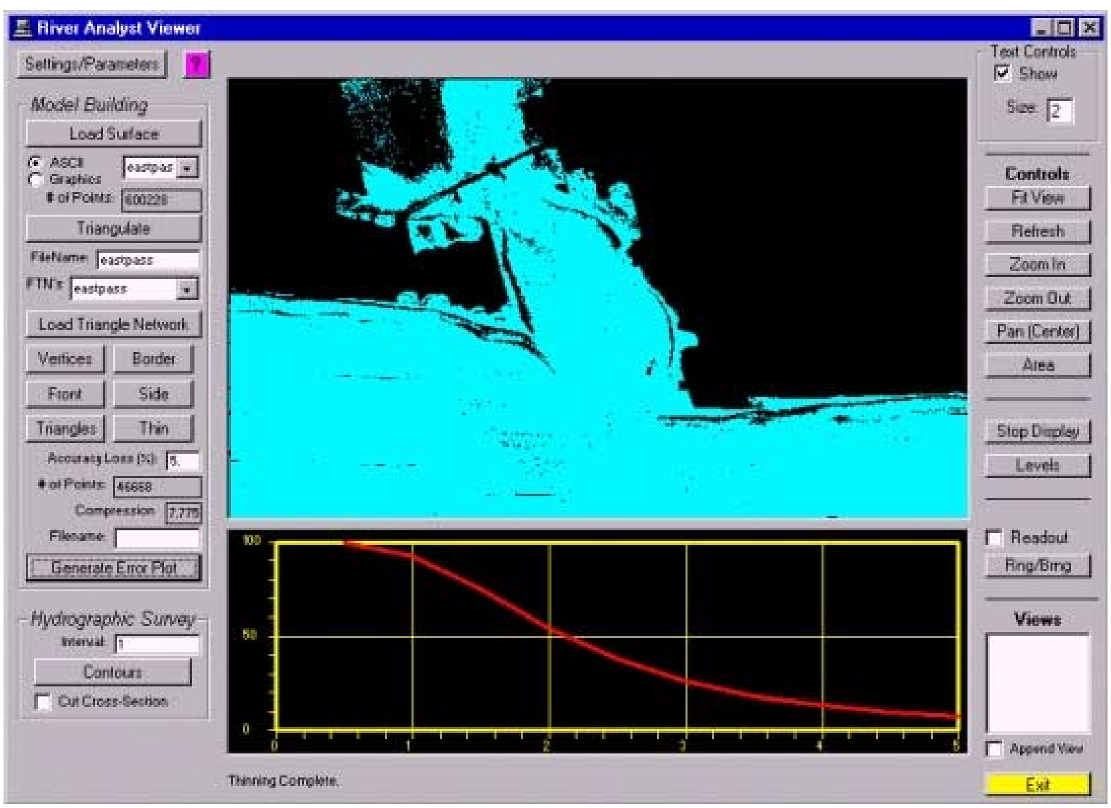
The Intelligent Thinning Algorithm

The intelligent thinning algorithm works by first fast-triangulating the entire set of easting, northing, elevation points into a Triangular Irregular Network (TIN). The traditional approach for generating TIN's has been to apply a Delaunay triangulation algorithm to the set of points. This ensures that no circle encompassed by the vertices of any triangle in the network contains any other points in the network. This is a very optimal method for modeling irregular surfaces, but it is extremely time-consuming to generate, particularly as the number of points in the surface increases. Beacon contends that triangle networks equivalent in quality to Delaunay networks can be quickly generated given that the vertices in the network consist of high density psuedo-randomly scattered points, as is the case in swath data. In fact, our Fast Triangulation algorithm processing time increases linearly with an increasing number of points while the Delaunay algorithm processing time increases exponentially. This Fast Triangulation first orders the collected points by sorting the Northing component of the position and then doing a subsort on Easting. Then the triangulation works by forming a moving "wall" of triangles. The first three points form a triangle and a wall consisting of two edges. Then, as points are removed, in order, from the sorted array of points they either attach within the extent of the wall, or extend the wall. As each point attaches to the wall, it forms a triangle and the wall advances in a Northerly direction until all of the points have been triangulated. The reason that this algorithm is so fast is that at any given time in the process, the computer has a minimum of information to keep track of. When a point is added in Delaunay triangulation, the computer must keep track of not only that point but of all of the other points in all of the other triangles which have been previously triangulated. That is why it slows down as the number of points increases. In Fast Triangulation, the computer only has to keep track of the point being added and the forward edge of the triangle wall. Anything previously triangulated behind the wall can be ignored. This algorithm has had great success with LIDAR data, multibeam survey data, and other high density data sets.

The intelligent thinning algorithm works by utilizing hypothetical contours. For an extremely dense set of contours, the thinning is performed by processing every vertex in the original triangle mesh and identifying those vertices which are NOT responsible for a single contour crossing in any of the triangles containing the vertex. Specifically, the thinning algorithm first marks every point in the vertex array with an ‘Unnecessary’ flag. Then, all the triangles are processed. This means that each vertex will be tested upon several occasions for different triangles. Each side of each triangle is processed. If the two vertices on a triangle side result in a contour threshold crossing, then BOTH vertices are marked as ‘Necessary’. Except for initially, vertices are NEVER marked as ‘Unnecessary’. After all sides of all triangles have been evaluated, the vertex array is rewritten containing only those points marked as ‘Necessary’. This technique ensures that areas with varying depths are not thinned while point density in relatively flat areas are greatly reduced. The contouring interval used is what controls the point reduction and also the accuracy loss.

Software Program Workflow

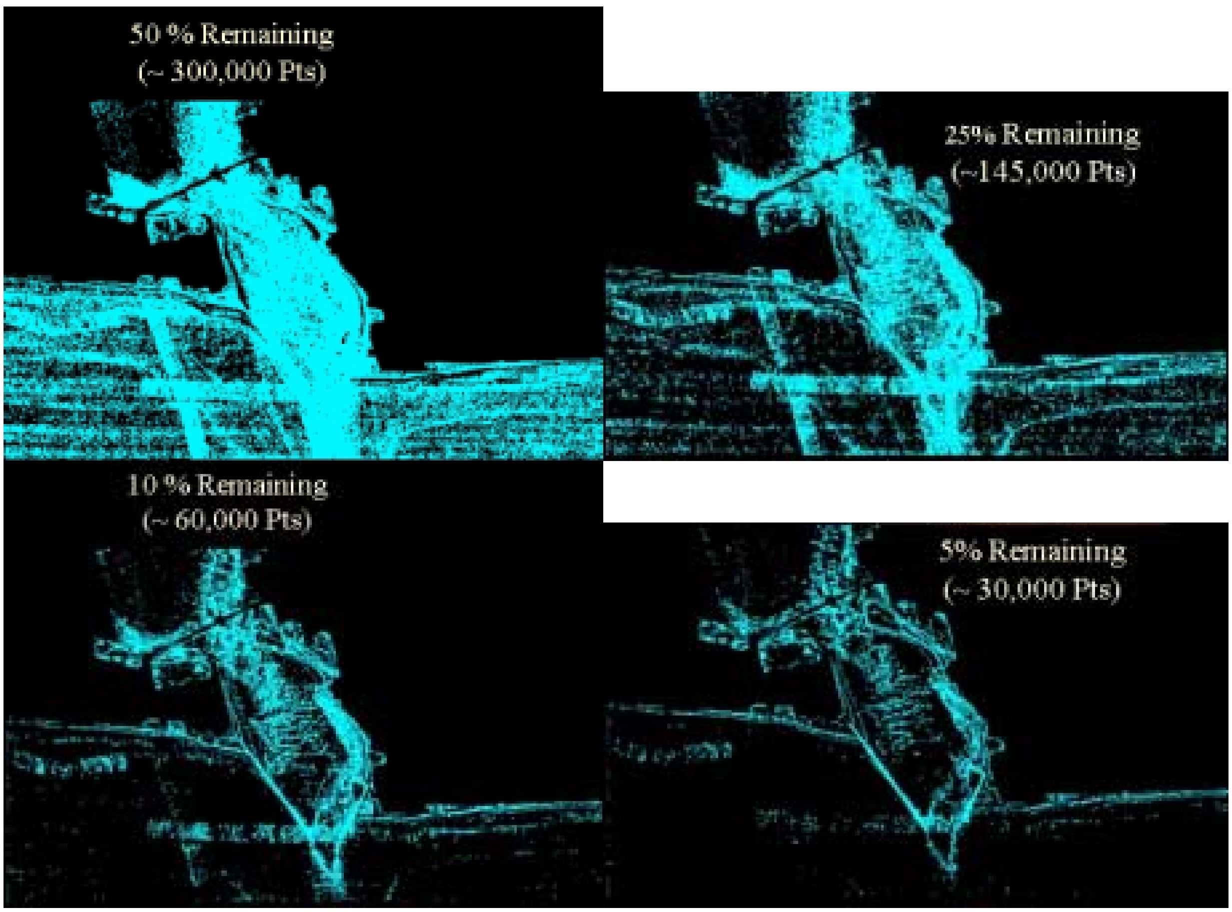
The IntelliThin sotware provides a very friendly user interface and straight-forward workflow. The operator is kept in the loop throughout the entire process. Capabilities are provided for initially reviewing the point data in top, side, and front views to evaluate initial data quality. Fence tools are provided for deleting obvious data points which lie outside of the main body of the data. During the data thinning process the IntelliThin software quickly thins the data at iteratively greater contouring intervals and presents a plot of point reduction versus accuracy loss to the operator who then specifies an optimal contouring interval and resultant level of thinning. This error plot allows an operator to visually judge the point of diminishing returns where increased accuracy loss does not result in significant point reduction. After thinning, additional software tools are provided for evaluating the quality of the thinned data set, including tools for a planar view of the thinned data points, contours, and cross-sectional profiles. The IntelliThin User Interface is shown below.



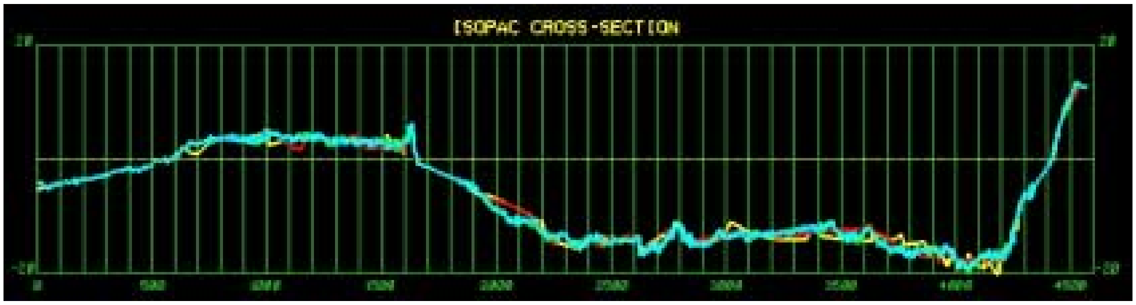
The unthinned vertices are shown in the planar view at the top of the interface form. The error plot, showing the tradeoff between amount of thinning and loss of accuracy is shown at the bottom of the form.

Algorithm Performance for a Sample Data Set

A sample data set for East Pass, Florida was utilized to quantify the intelligent thinning algorithm performance. For the baseline xyz ASCII file of 600,228 points the triangulation time was 14 minutes on a 225 MHz Pentium with 128 Mbytes of RAM (on an 800 MHz Pentium III, the triangulation time was only 4 minutes 45 seconds). Once triangulated, the program only required 1 minute to perform thinning at a specific accuracy loss. The program automatically made 10 passes from .5% to 5% accuracy loss. This corresponds to 0 to 95% reduction in the number of points in the data set. The following figures illustrate the points kept at representative point reduction factors of 50%, 75%, 90%, and 95% thinning.



The error plot showed that 3% accuracy loss thinning was the ideal point for trading off accuracy for point reduction. So, with one more pass to actually thin at 3% accuracy loss or approximately 75% reduction in points (which directly corresponds to 2.9 foot contours for this data) the total time required for thinning was 25 minutes. After thinning, there were 156,572 points (26% remaining). An isopach comparison of the the reduced triangulated data set compared to the original triangulated data set showed a maximum difference between the surfaces of only 3 feet. Cross-sections were also cut across each of triangulated surface corresponding to each of the representative point reduction factors. This cross-section plot is shown on the following page.



This cross-section shows minimal differences between the triangulated surface models, indicating that the intelligent thinning algorithm did indeed ensure that areas with varying depths are not thinned while point density in relatively flat areas is greatly reduced.

Summary

The Intelligent Thinning software, IntelliThin, has several distinguishing features. An innovative fast non-Delaunay algorithm which generates "Delaunay quality" triangles is utilized for performing an initial triangulation of the high density data set. This initial triangulation is the only triangulation required regardless of the number of data thinning passes. An intuitive error plot lets the user judge the trade-off between reduction in points and accuracy loss. This trade-off point represents the point of diminishing returns where additional loss of accuracy really does not result in that much more reduction in points. Finally, the software features an intelligent thinning algorithm which retains points where the elevation is changing and discards points which only minimally affect the triangulated terrain model.

Author Biography

Michael Grounds has been specializing in graphics applications for River Engineering for the past ten years working primarily with the Memphis and Vicksburg Districts of the Corps of Engineers. He has developed programs which apply to the entire lifecycle of Revetment Design, Construction, and As-Builting. He also developed the analysis programs used in Memphis for micro-modeling and in St. Louis by the Applied River Engineering Center for Doppler Data Analysis. He has developed the Corps-wide Survey Engineering Monumentation Management System, www.semms.org, and is currently involved in the development of a nationwide Dam Safety Program Performance Measures Database, www.safedams.org. Mr. Grounds currently specializes in real-time GPS based applications for electronic navigation, hydrographic surveying, dredging, channel patrol, and geophysical surveying. These programs are currently in use in the New Orleans, St. Louis, Memphis, and Vicksburg Districts. Mr. Grounds has been employed for 23 years at Intergraph Corporation, and has been employed for 4 years at Beacon Resources.