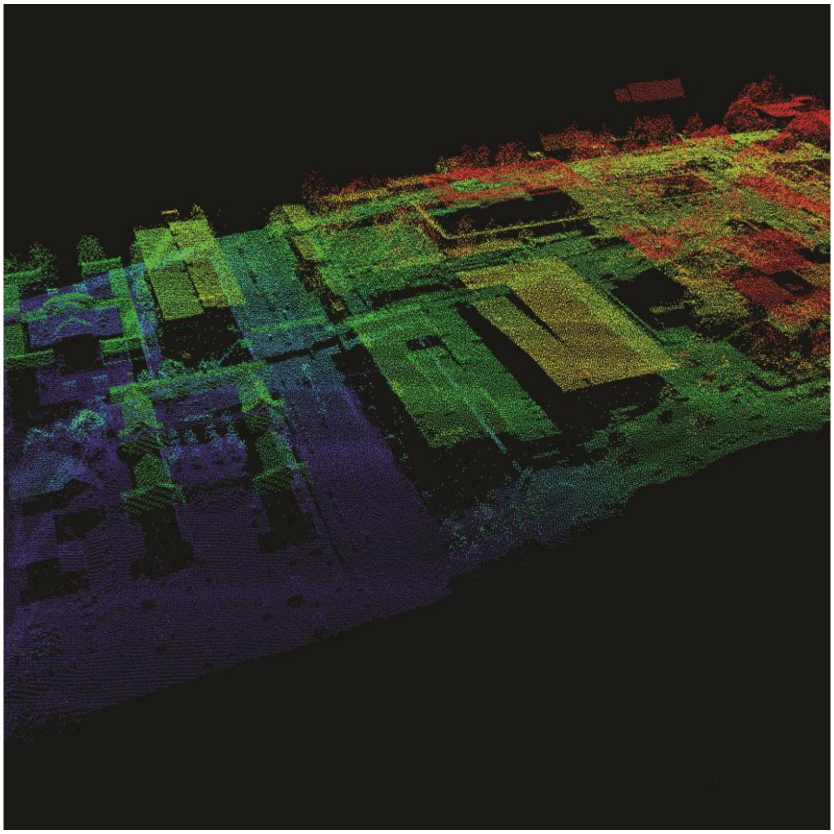
REIN IN 3D POINT CLOUDS WITH THE LAS FORMAT

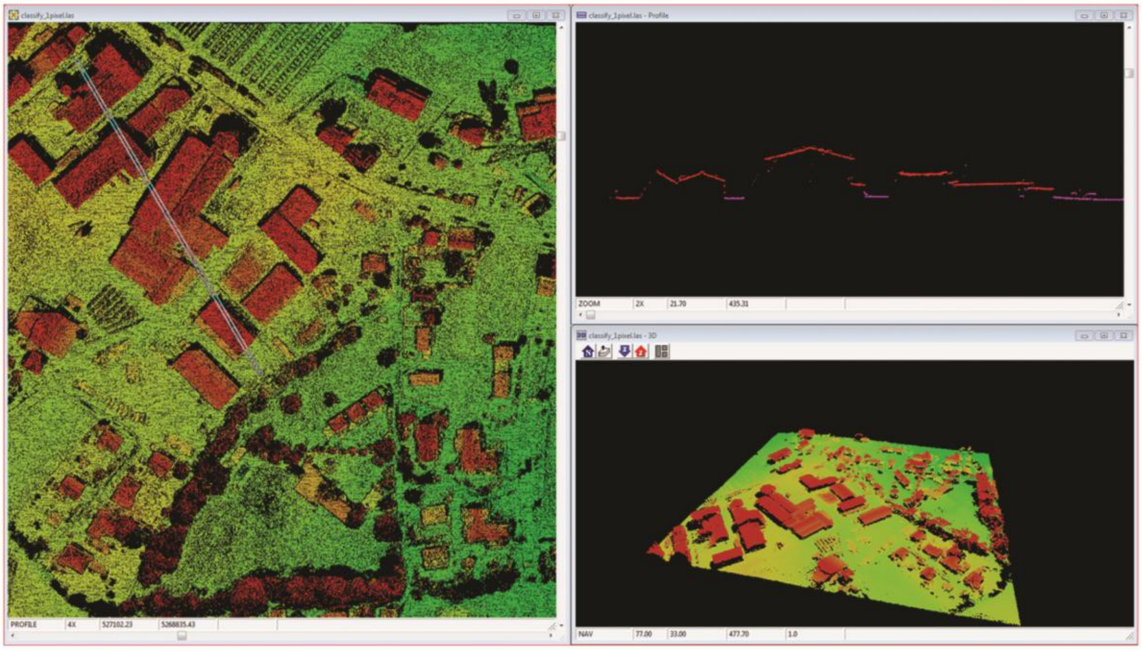
BY AMBER CHAMBERS AND DR. TOM LOBONC

The advent of LiDAR sensors has been valuable for the geospatial industry, but there have been some challenges as the community attempts to fully leverage the technology. As you may already know, LiDAR is similar to radar in that both direct an active signal towards the surface of interest and receive a return signal reflected from the surface and objects on it. The time delay of this reflection is used to calculate range from the sensor to the surface. With radar, this results in formation of an “image” with intensities proportional to the reflectivity of the surface at radar wavelengths, which are in the microwave or radio range. In contrast, LiDAR processing is used along with sensor position information to determine the elevation of the surface at each reflection point, or the distance from an established reference plane in the case of terrestrial LIDAR.



Points from a 3D LiDAR point cloud, color-coded to show relative elevation.

As a general rule, you can only detect things that are at least as large as the wavelength being directed at them; therefore, LiDAR, with its shorter wavelengths in the ultraviolet to near-infrared range, is capable of detecting smaller objects than radar with its relatively longer wavelengths. Additionally, more objects are reflective (and many more strongly reflective) in the LiDAR wavelengths than the radar wavelengths and since infrared’s shorter wavelength means a higher frequency, LiDAR is able to send out more pulses in a given period of time than radar, allowing data to be collected on more locations within the scanned area. LiDAR also can register multiple returns from the same signal pulse when that pulse is reflected from multiple different heights. For example, a single pulse may reflect partially from a leaf at the top of the tree, a branch partway down, and the ground under the tree. LiDAR data vendors then use the raw return data along with sensor position and orientation information from the GPS and IMU to calculate the x, y, and z coordinates of multiple points in the scanned area, resulting in a dense 3-dimensional point cloud. For each point in the cloud, the intensity of the return is also persisted.



Left: Points in a 3D cloud generated from stereo imagery; Top right: Profile from the 3D Point Cloud; Bottom right: 3D model generated from the point cloud.

These intensity values can be used to visualize the point cloud so that it appears similar to a panchromatic image. Details about the sensor itself and the flight information are often useful and documented to accompany the point data. In scientific endeavors, more information is usually a good thing, and this has been the case with LiDAR; it just presents the issue of effectively organizing and storing the data so it can be used most efficiently.

Storing LiDAR Data Efficiently

Originally, data vendors and software companies often used different proprietary formats to store the data, and the format supplied by a vendor might not be compatible with all of the software packages used by potential customers. The formats required by the software could also vary between applications, making it difficult for organizations using different software to collaborate. An alternative to the proprietary formats is to store the data in an ASCII text file with the information written as a series of comma-separated values. ASCII files are a little more interoperable than the proprietary formats, but still aren't perfect in this area unless there is some type of agreement as to how the information will be presented within the file. Due to the detail of most LiDAR data, these files can become very large, and consequently, slow to open, even when storing only a relatively small amount of information.

The LASer (LAS) format was first developed by EnerQuest as its own proprietary method of storing LiDAR data. LAS is a binary format that stores LiDAR metadata and all the information about a 3D point cloud in a relatively compact form. It was also designed so that all of the original data can be retaine. The LAS format allows each point in the cloud to be assigned a class, such as ground, low vegetation, high vegetation, or building. This enables the creation of multiple value-added products from the same dataset without altering it. For example, software utilizing an LAS file could suppress points classified as vegetation or buildings to focus solely on the ground points for the creation of a bare earth digital elevation model (DEM or DTM). Or, the same file with all the same data could be used to create

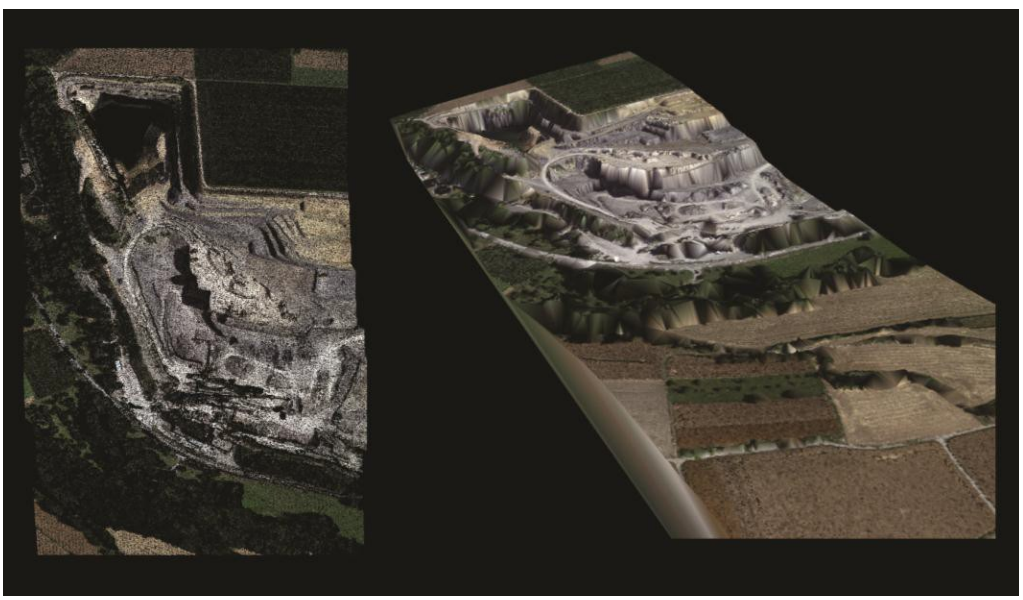
3-dimensional digital surface models (DSM) of the above-ground structures along with visible ground. Also, due to the sensitivity of LiDAR, it will detect objects such as birds and atmospheric aerosols and dust. The format permits these points to be marked as suspected noise so they can be withheld when products are being generated, but they do not have to be deleted from the file. Finally, LiDAR point clouds can be fused with optical imagery (typically when collected simultaneously with the LiDAR) so that the image pixel RGB values are also encoded into the LAS file. This allows for an instant photorealistic 3D visualization.

In 2002, EnerQuest contributed the LAS format to the public domain and industry stakeholders such as EnerQuest, Z/I Imaging, Optech and Leica Geosystems began the process of standardization. The American Society for Photogrammetry and Remote Sensing (ASPRS) published the ASPRS LiDAR Data Exchange Format Standard, version 1.0 in May 2003.

Keeping Up With the Industry

Initially, the LAS format was intended to be user-extensible, but user-defined extensions can hamper the goal of interoperability. For example, with LAS 1.0 points could be classified using a number that represented a class, but there were no standards for commonly-used classes and the ground might be class 1 on one file and class 2 on another. Three more versions of the standard have been developed to address both technological updates and the various needs of the user community so that it can be adhered to universally and further the original goal.

There have been several noteworthy advances during this sequence of updates. In 2005, Version 1.1 introduced the complete standardization of classification by assigning a specific number to commonly used classes such as ground, low vegetation, medium vegetation, and buildings. Version 1.2 (2008) brought the ability to store RGB color values for a point, which can be fused in processing with elevation values to increase classification accuracy, as well as providing the photorealistic 3D models mentioned earlier. The most recent update, version 1.3 (July 2009) added support for storing full waveforms of the entire backscattered laser pulse rather than just the elevation return values. Full waveforms can provide more detail and information about the structure of scanned objects through analysis of the waveform shape.



Left: A point cloud shown with RGB values for each point;
Right: An RGB TIN generated from the point cloud

Starting in version 1.1, the LAS format specification has allowed users to specify the software that generated the file, or that a particular point is synthetic instead of originating from LiDAR sensor data. These metadata attributes are particularly useful for instances where the data doesn't come from a LiDAR sensor, which is more frequently the case now. Obtaining LiDAR data can be expensive, since LiDAR coverage from a given flight line is much less than that of a large format aerial camera, requiring more flying time and greater post-processing time. Due to the cost of LiDAR data and the increasing number of sensors already delivering vast amounts of high-resolution stereo imagery every day, there has been increasing demand for new software applications that use algorithms to create dense 3D point clouds from stereo imagery. These generated point clouds can be saved in LAS format the same as LiDAR data, can be readily encoded with RGB values, and can be used to produce most of the same value-added products, such as digital elevation models that can be used in orthophoto production, 3D surface models, and information products such as viewshed and mobility analysis. These algorithms are very sophisticated, using such techniques as pixel-wise matching, in which every pixel in the source imagery is utilized to obtain points in the cloud, segmentation to constrain processing and assist in object identification, adaptable algorithm parameters, and automated blunder detection. Additionally the most sophisticated algorithms offer the ability to use imagery and the generated elevation information to assign object classifications for each point, which are then encoded in the output LAS file. Software using these advanced techniques has been proven capable of accuracy as good as, and in some cases, slightly better than actual LiDAR data.

LiDAR has rapidly grown from a niche technology to one that is in common use in both commercial and military surface generation and for information extraction. The technology is advancing at a rapid pace and new applications and improved information extraction algorithms are appearing regularly. Keeping pace with the technological developments, LAS has now become the de facto standard for storing point cloud data acquired from LiDAR (and other sources). As the first real solution featuring interoperability, LAS also provides compact and efficient data storage and flexibility. In addition, LAS continues to have broad community support working to keep the format current.

Amber Chambers is Technical Marketing Specialist at Intergraph and has a background in software engineering and business administration.

Dr. Tom Lobonc is Director of Defense Products at Intergraph and has a background in geospatial production systems and workflows.

About Hexagon Geospatial

Hexagon Geospatial helps you make sense of the dynamically changing world. Known globally as a maker of leading-edge technology, we enable our customers to easily transform their data into actionable information, shortening the lifecycle from the moment of change to action. Hexagon Geospatial provides the software products and platforms to a large variety of customers through direct sales, channel partners, and Hexagon businesses, including the underlying geospatial technology to drive Intergraph® Security, Government & Infrastructure (SG&I) industry solutions. Hexagon Geospatial is a division of Intergraph® Corporation. For more information, visit www.hexagongeospatial.com.

Intergraph® Corporation is part of Hexagon (Nordic exchange: HEXA B). Hexagon is a leading global provider of design, measurement and visualisation technologies that enable customers to design, measure and position objects, and process and present data.

Learn more at www.hexagon.com.