PE&RS December 2002

VOLUME 68, NUMBER 12
PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING
JOURNAL OF THE AMERICAN SOCIETY FOR PHOTOGRAMMETRY AND REMOTE SENSING

Highlight Article

Product Definitions and Guidelines for use in Specifying Lidar Deliverables

Martin Flood

Since the mid-1990s the science of laser altimetry has been actively adopted by the remote sensing community as a tool for the rapid generation of accurate, high-resolution, digital terrain models. The fundamental science behind laser altimetry has been studied for decades, going back to the 1960s, but it is only with the increased availability of off-the-shelf sensors and the emergence of commercial data providers that this tool has seen wider deployment and acceptance by geospatial data users. The advantages of laser altimetry include rapid turn-around times, generation of relatively high-accuracy, high-density data sets and the ability to map in areas of low contrast, low relief or relatively dense vegetation cover. The disadvantages include relatively high cost on a project basis, a lack of direct object-oriented information (imagery, spectral information), data processing issues related to robust, efficient feature extraction (bare earth, break lines) and a lack of common standards and professional practices. As a result laser altimetry, or lidar mapping as it is more commonly referred to in the mapping community, is increasingly used in conjunction with other sensors ranging from traditional film imagery to integrated digital cameras, synthetic aperture radar or hyperspectral scanners.

While lidar mapping is now available from a variety of commercial data providers as well as various academic and research groups, most lidar data users are not very familiar with the technology. This situation has caused problems in the past when different expectations and assumptions between the data provider and the data user about the capabilities of lidar mapping and the data products to be delivered have led to failed projects or disappointing results. The situation has improved over the past two years as data users have become more experienced with the technology and data providers have become more conservative in their promotion of lidar's capabilities, but many organizations are only just starting to consider lidar-derived elevation data as a routine part of their mapping programs. Guidelines for the acquisition, analysis, manipulation and implementation of lidar-derived elevation data are needed to further educate the geospatial community about the advantages and disadvantages of using lidar mapping. The ASPRS Lidar Committee, part of the Photogrammetric Applications Division, is working on a Handbook of Operational Lidar Mapping to address these issues. This is intended to be a comprehensive reference on best practices and recommended guidelines for the professional implementation of lidar mapping. The Handbook is planned for release in 2003. However there are some commonly accepted product definitions and guidelines already in use that a contracting agency can use when considering acquiring lidar data from a commercial data provider to help ensure a successful project. This article provides a summary of these product definitions as a reference and as the basis for common discussion between contracting agencies and data providers.

Lidar Data Product Definitions

As a relatively new technology, lidar mapping has yet to develop a commonly accepted set of data product definitions. As a result a data user may be presented with different deliverables when dealing with different data providers, which can make it difficult to compare competing bids or to estimate the additional work that will be required to integrate the lidar-derived elevation data in to the mapping project. It is important to ensure an appropriate set of deliverables will be received from the data provider for the mapping project at hand. The selection of the optimum product to support additional work such as orthorectification or contour generation is an important component of a successful mapping project. At present there are some standard data products that are generally accepted as “common” deliverables and any organization contracting for lidar data should be familiar with these products.

Level 1 - Basic or “All-Points”

The basic lidar data set is the “all-points” or “all-shots” point cloud. This is essentially all of the post-processed lidar data properly geo-referenced but with no additional filtering or analysis. Even for small projects this can easily run in to the tens of millions of individually geo-referenced data points. Until recently this massive point cloud was not considered a standard deliverable as few clients had the capabilities to view, let alone extract information from the point cloud. [See Figure 1. http://www.enerquest.com/silc2.htm Unfiltered lidar colored by elevation.] The software tools for doing so were proprietary to the data collectors or instrument manufacturers and most data users had to develop their own classification and analysis tools or request these services from their data provider. However, the emergence of several off-the-shelf software packages for the analysis and manipulation of large point cloud data sets as well as several public domain algorithms has allowed many organizations to develop this expertise in-house, thus allowing them to outsource just the lidar data collection portion of a project if desired. This in-house approach has also become more practical as the available pool of trained and experienced lidar data analysts has increased. Various levels of skilled personnel can now be hired relatively easily and normal staff turnover and the movement of human resources between organizations has helped to diffuse best practices and common data handling techniques across the industry. For any organization that intends to routinely use lidar point cloud data in its activities and has the resources to develop the analysis, classification and manipulation capabilities in-house, the “all-shots” product is the desired deliverable.

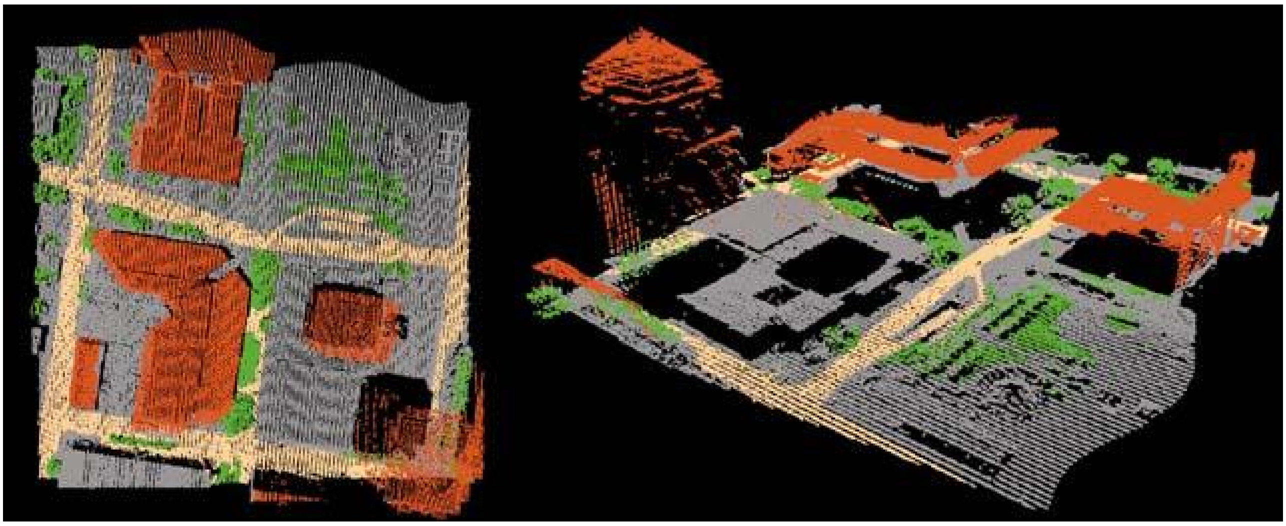

Figure 1. Lidar returns classified by features, roads, surface, buildings and vegetation, using SILC technology

The increasing demand for the basic all-shots product is also being driven by the emergence of organizations offering dedicated point cloud processing using either proprietary or third-party tool sets. By consolidating demand across numerous projects and optimizing processing facilities for handling point cloud data, these service bureaus offer a cost-effective alternative for organizations looking to avoid having the lidar data collector handle the processing but not interested in dedicating their own internal resources to lidar data analysis and manipulation. The adoption of a common binary format for the exchange and analysis of lidar data, the LAS industry standard format, will further increase the appeal of this service bureau approach, allowing contracting agencies to effectively partner with best-in-class firms on both the data collection and data processing portions of a project without being captive to a single organization for both.

Level 2 - Low Fidelity or “First-Pass”

The most common value-added product produced by lidar data providers at present is what is referred to as “first-pass” or “preliminary” classification and filtering of the lidar point cloud. Using either proprietary algorithms or third-party software tools, the data collector will automatically filter the point cloud in to points on the ground, the “bare earth,” and points that are not ground, such as vegetation, buildings or other man-made features. The resulting product is a low-fidelity terrain model that may still contain misclassified ground/non-ground points. These points will be delivered separated in to two layers; ground and non-ground (or vegetation). There is generally no classification of the non-ground points in to separate features types (buildings, trees, etc.) and the ground points generally include some percentage of residual features not extracted by the automated classification algorithms. [See Figure 2. http://www.enerquest.com/silc2.htm First pass filtered, colored by elevation.] For applications that do not require information about the above-ground features and for which a high fidelity terrain model is not a necessity, these ground/non-ground layers may be adequate as the final product. While a data user with the appropriate software tools can accomplish the same separation of the full point cloud in to ground and non-ground, it is often more efficient to have the data collector perform this step as it is essentially fully automated.

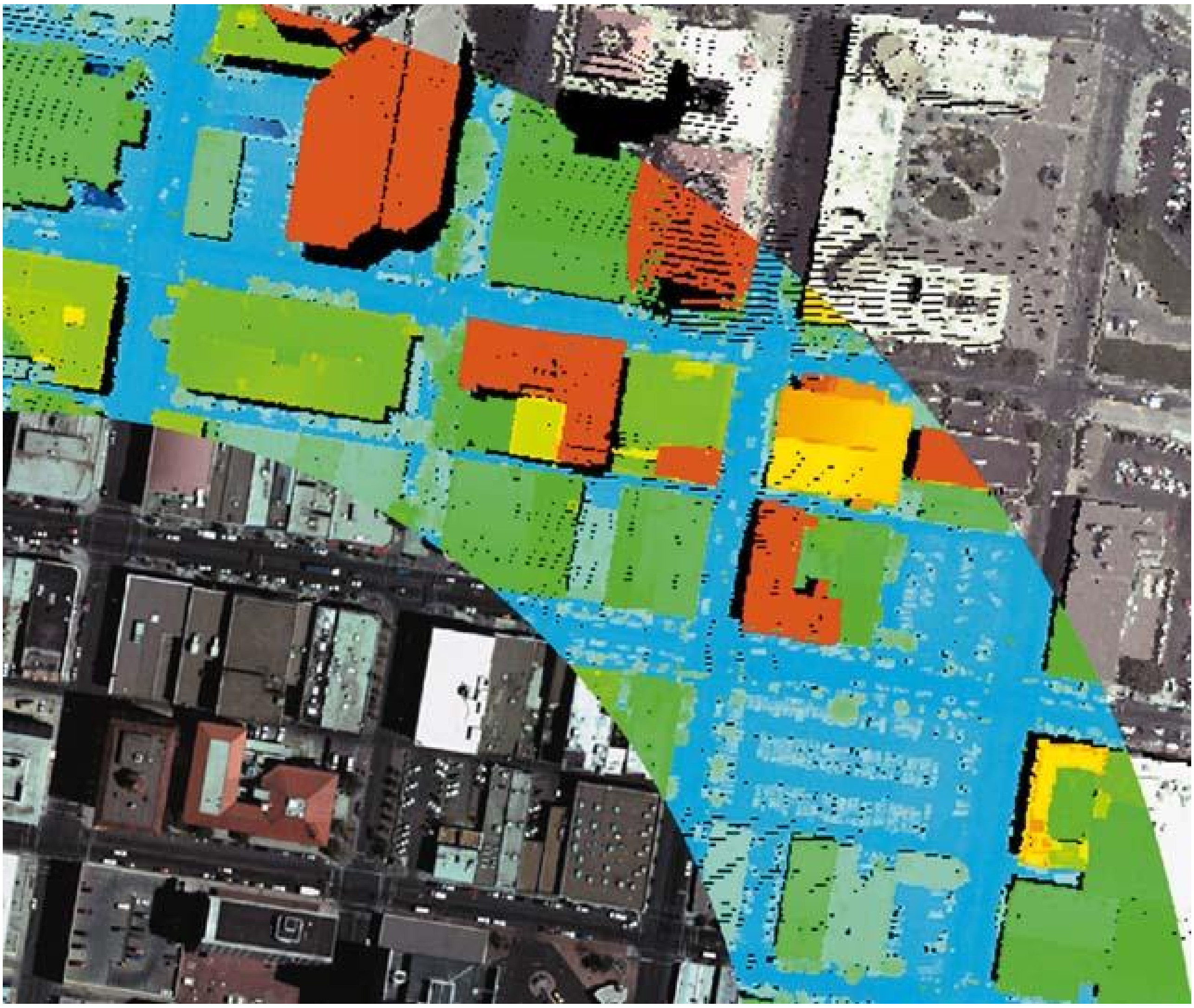

Figure 2. Inner circle – color digital imagery. Middle circle – Lidar data colored by elevation. Outer circle – SILC laser points attributed with SILC RGB values.

Significant effort is under way by various groups to develop more robust and efficient automated filtering techniques that improve the fidelity of the first-pass terrain model or improve the feature extraction capabilities of the automated algorithms. These approaches often include the integration of object information provided by the intensity of the laser pulse return, simultaneously acquired digital imagery or direct spectral tagging of the elevation data. It is likely that incremental improvements to automated classification will continue to be made over the next few years but that “100%” automatic classification from lidar elevation data alone will remain a goal rather than a routinely achievable specification. When requesting only a low fidelity terrain model based on preliminary classification as the deliverable product, it is important to specify how aggressively the data collector will set the automated classification routines. A more aggressive setting will provide a higher fidelity terrain model — a cleaner data set — but with an increased number of misclassifications, while a less aggressive setting will reduce misclassification but produce a lower fidelity terrain model. If the contractor is considering doing significant work in-house on the elevation data, it is advisable to specify a less aggressive first-pass approach. In all cases it is useful to request documentation of the filter settings (e.g. iteration angles, maximum distances) used for any third-party software classifications so that these can be repeated if necessary. The use of subjective terms such as “85% clean” should be avoided in any contractual specifications for preliminary or first-pass data products. Unfortunately, this type of requirement is becoming more common in RFPs requesting lidar data but the author is not aware of any robust, reliable method for accurately verifying that such a specification has been met by the data provider.

Level 3 - High Fidelity or "Cleaned"

While preliminary, first-pass automatic filtering is a necessary step in generating high fidelity terrain models from the lidar elevation data, it is usually not sufficient. In any area with more than a minimum of ground cover or man-made features or where the topography deviates from an open, flat plane, residual artifacts and misclassifications in the data set will require further analysis. Common problems remaining in a preliminary product include poor ground model fidelity in areas of low, dense ground cover, the inability to accurately capture sharp grade breaks such as low ridges or sharp cuts, misclassification of man-made features such as bridges, and an inability to discriminate tree cover from topography in areas of sharp relief. Benchmarks and quantitative reports of the efficiency and accuracy of automated classification routines are generally not published but it is commonly quoted that most automated classification routines are 80% to 90% effective. Depending on the local ground cover and topography, they will accurately classify 80% to 90% of the ground points or remove 80% to 90% of the non-ground points. The remaining 10% to 20% of the data needs to be analyzed and classified manually either with supporting imagery or directly by a trained lidar data analyst.

If requested, most lidar data providers will deliver fully edited data sets that have been extensively reviewed by an experienced data analyst to remove any artifacts created by the automatic classification routine and provide a "99%" clean terrain model. [See Figure 3 http://www.enerquest.com/silc2.htm Bare Earth "cleaned" lidar colored by elevation.] Again, the 99% claim is often a subjective evaluation not a rigorously verified specification. However, this process, even if aided by accompanying imagery or similar data, is a labor-intensive step that will add to the project costs and the schedule. As project size increases, the ability of even highly skilled lidar data analysts to match the throughput of airborne data collection and initial automated classification rapidly decreases resulting in a bottleneck at this manual editing step. As a result, any contracting agency considering large lidar data collection programs, such as statewide efforts covering thousands of square miles, should pay particular attention to staffing requirements and schedule impacts imposed by the need for manual classification of the data. Estimates of the required staff and the anticipated throughput from manual review and editing should always be requested from the lidar data provider. Underestimating the necessary effort to complete these activities puts schedule and budget performance at risk. Smaller projects are less impacted but contracting agencies may still want to consider doing final manual review and editing in-house to avoid duplication of effort with the data collector.

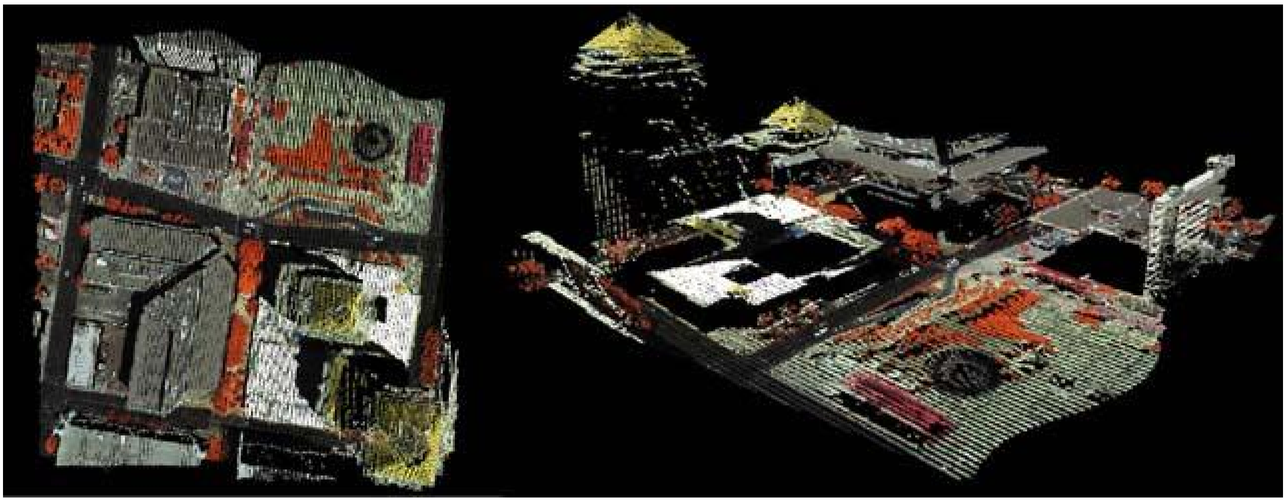

Figure 3. Unfiltered Lidar data colored with SILC (CIR values), plan view and oblique.

Level 4 - Feature Extraction

Feature extraction is the next stage of value-added lidar data processing that is used to generate application-specific data products. Once the high fidelity terrain model has been extracted from the full point cloud, the remaining data contain information about the above ground features in the project area. These can be natural vegetation, man-made features or a combination of both. These features can be of interest to the data user or they can be essentially viewed as noise to be discarded. Depending on the application, further analysis of the non ground points can be completed using another combination of automated and manual classification to identify features of interest. For example, the extraction of power lines strung between utility towers allows the accurate calculation of the catenary curve of the wire, information of value to power utilities; creating a model of the canopy surface to allow canopy height modeling is of interest to foresters; extracting building footprints and rooftops from city models is of interest to various groups working in urban environments. [See Figure 4 opposite page & http://www.enerquest.com/silc2.htm Lidar return classified by feature using EnerQuest SILC (Spectral Imagery Lidar Composite) process.] In general this type of value-added feature extraction is handled by application-specific tools developed independently of the lidar data collectors or instrument manufacturers. Some of the more common tools, such as power line extraction, are being integrated in to common third-party software tools. If the project requires application-specific deliverables such as power lines or building footprints, it is important to specify these explicitly in the contract and determine the capabilities and experience of the data collector in this specific area. Again, using a service bureau or third-party data processor that specializes in the desired application and has experience or has developed customized tools for the specific type of feature extraction can be the most cost-effective solution.

Table 1. Product Definitions for Lidar Data

Level	Name	Description
1	Basic or "All Points"	All of the post-processed lidar data properly geo-referenced but with no additional filtering or analysis. Suitable for those organizations with in-house data processing tools and capabilities or who work with a third-party data processing service bureau. Cheapest and fastest product.
2	Low Fidelity or "First Pass"	Using either proprietary algorithms or third-party software tools, the data provider will automatically filter the point cloud in to points on the ground, the "bare earth", and points that are not ground. There is generally no classification of the non-ground points in to separate features types (buildings, trees, etc.) and the ground points generally include some percentage of residual features not extracted by the automated classification algorithms. Suitable for those organizations with in-house data processing tools and capabilities or who work with a third-part data processing service bureau. Common deliverable. Usually same cost/schedule as All-Points
3	High Fidelity or "Cleaned"	A fully edited data set that has been extensively reviewed by an experienced data analyst to remove any artifacts created by the automatic classification routine and provide a "99%" clean terrain model. The low fidelity data are analyzed and classified manually, usually with supporting imagery. Labor-intensive product. Moderate cost but with longer delivery schedules, especially on larger projects.
4	Feature Layers	Further processing using a combination of automated and manual classification to identify features of interest such as power lines or building footprints. Generally completed in-house or using a service bureau or third-party data processor that specializes in the desired application and has experience or has developed customized tools for the specific type of feature extraction. Usually more expensive product than high fidelity terrain model.
5	Fused	A further refinement of the lidar data product achieved by the fusion of the lidar-derived elevation data set with information from other sensors. This can include digital imagery, hyperspectral data, thermal imagery, planimetric data or similar data sources. Generally the most information-rich product with the highest cost.

Level 5 - Fused

A further refinement of the lidar data product can be achieved by the fusion of the lidar-derived elevation data set with information from other sensors. This can include digital imagery, hyperspectral data, thermal imagery, planimetric data or similar data sources. The additional information may have been captured simultaneously with the lidar data or collected separately during the field campaign or even procured at a different time. Again the cost/benefit of having this fusion done by the lidar data collector needs to be weighed by the in-house capabilities of the data user or by the availability of a third-party service bureau to handle the integration. Fused data sets are the most information-rich data products created from lidar-derived elevation data. [See Figures 5 & 6 (previous page & http://www.enerquest.com/silc2.htm) Lidar and natural color (RGB) fused using the EnerQuest SILC process, lidar and CIR imagery.]

Summary

The use of elevation data derived from lidar mapping can greatly enhance the resolution, accuracy and cost-effectiveness of a mapping product. However, it is important to match the requested lidar data product to the specific needs of the mapping project. Over-specification, such as requesting intensity data when it will not be used, can unnecessarily increase the cost of a project. An analysis of various alternate approaches to acquiring the final data product from a single source, for example working with both a data collector and a data processor or developing in-house data processing capabilities, may identify cost savings for organizations on a project-by-project basis. By using a commonly accepted set of product definitions and clearly and accurately specifying the required lidar data products, mapping agencies should be able to streamline their acquisition process and reduce operational risks due to differing expectations or inappropriate data products.

Author

Martin Flood is an independent industry analyst specializing in the commercialization of laser-based remote sensing technology. He has worked with numerous organizations in the commercial lidar industry and is the original creator of the airbornelasermapping.com web site. He currently serves as the Chair of the ASPRS Lidar Committee, part of the Photogrammetric Applications Division of ASPRS. He can be reached at flood@lidarcentral.com.
All Images courtesy of EnerQuest Systems LLC

For more information and examples on SILC process see http:/www.enerquest.com or contact Richard A. Vincent at rvincent@enerquest.com

---

Copyright© 2002 American Society for Photogrammetry & Remote Sensing
5410 Grosvenor Lane, Suite 210, Bethesda, Maryland 20814-2160

Page URL: http://www.asprs.org/asprs/publications/pe&rs/2002journal/december/highlight.html

http://www.asprs.org, for assistance select here.
(11/25/2002 18:58:09)