Mapping Matters

Your Questions Answered

The layman's perspective on technical theory and practical applications of mapping and GIS

By Qassim A. Abdullah, Ph.D., PLS, CP**

Can you clarify waveform or signal echo digitizing? How is this data provided and how would you analyze it? Is the echo signal return intensity? Phase?

submitted by a participant of the ASPRS webinar "Lidar Fundamentals and Applications"

Dr. Abdullah: laser waveform has been used in the research community for a while. However, the remote sensing and geospatial community started hearing about it only in the last few years. Conventional topographic Light Detection and Ranging (lidar) is characterized by sending and then receiving the diffused/reflected laser signal as discrete laser shots that may count up to 4-5 returns for each emitted pulse. The system uses the time it takes the pulse to travel to the ground feature and then bounce back to the sensor in conjunction with the speed of the laser pulse, which is the speed of light, to extract the range (distance) between the sensor and the feature. By combining the measured range of the laser with data from GPS and Inertial Measurement Unit (IMU) for each return, the system establishes 3-dimensional coordinate locations at the point where the laser signal is bounced back. These locations in a 3-dimentional space form the digital surface model or point cloud that represents the ground and what is covering it. That is what I called the "low hanging fruit" era of the laser applications for geospatial mapping activities from which we reaped its benefits for the last two decades in the form of conventional or discrete lidar. The low hanging fruit era, which is characterized by detecting the arrival of the returned pulse (peak), is based on the simple thresholding technique where the return(s) of a laser pulse is recorded only if the strength of the returned pulse(s) exceeds certain threshold. Discrete lidar systems, as used in the last two decades, can collect up to 5 returns from a single emitted pulse. However, the user rarely finds valuable data, if any, recorded after the third return.

> *"Most new systems are flexible in terms of giving the user the options of collecting discrete data alone or discrete data and waveform making them more suitable for the task at hand."*

Advancement in technologies over the last decade provided newer lidar systems with the capability of recording signals of the entire backscatter (here I use the terms backscatter, echo, and return to refer to the diffused and reflected laser pulse throughout my answer) in contrast with the discrete lidar, which records only a few returns. Almost all existing lidar systems available in the market today are modified one way or another to enable it to collect waveform data. Some of these systems are upgraded by adding a Full Waveform Digitization (FWD) module to the controller of the existing discrete lidar. Manufacturers also boosted the storage capacity of the system to prepare them for full waveform digitization capability.

The new module added capability to the current discrete lidar system by sampling and recording waveform data (amplitudes) of the returned signal from a laser pulse fired by the system. A FWD-enabled system records both discrete returned pulses and additional data to define its waveform. In current discrete lidar, the received returned pulse is split and transmitted to the range finding circuitry of the discrete lidar (resulting in 3-D point cloud) and to the intensity digitizer. In FWD-enabled lidar, the signal that used to be sent to the intensity digitizer is now directed to the waveform digitizer where it is split and passed to the intensity digitizer. The recorded waveform data is precisely referenced by means of the time synchronization unit to that specific discrete pulse return from which the waveform data is collected and recorded.

Most new systems are flexible in terms of giving the user the options of collecting discrete data alone or discrete data and waveform, making them more suitable for the task at hand. In addition, new systems allow the user to control the characteristics of the wave form (longer or shorter waveform) by adjusting the time resolution. In other words (and in addition to the typical point cloud and intensity image), the new FWD lidar systems continuously record the entire profile of the reflected energy of the emitted pulse over a span of time at pre-selected time intervals (usually 1 to 10 nanoseconds) This results in a uniform sampling of the object depth along the laser path. The process of temporal slicing of the signal provides a wealth of information about the ground object that cannot be described by the few returns recorded by the conventional discrete laser systems.

> *"The waveform is more than just intensity, as current intensity is based on thresholding technique. Intensity lacks the intricate details that the indefinite digitizing capability of the waveform digitizing system provides about the returned signal or echo from objects along the laser path."*

continued on page 106

continued from page 105

In addition, the waveform can also be used to extract additional 3-D discrete points that can be added to the points routinely collected by the system. This results in a denser Digital Surface Model (DSM) than the one obtained from discrete-mode operation alone.

To understand the principal of FWD system, consider it like CATSCAN imaging. Rather than having merely an exterior image of the organ, it provides the physician with clear digital images of systematic slices through the patient's organ at different depths. Although the new technique is referred to as "full waveform digitization", some available lidar systems today record part of the waveform over certain periods of time. Most systems limit the maximum number of samples to be recorded per pulse to a few hundred. The user-defined sampling rate (or interval) of the waveform determines the distance resolution (or vertical resolution) while the sample depth (e.g., how far the signal travels through the forest canopy) is determined by the maximum duration of recording.

> *"In addition to the typical point cloud and intensity image, the new FWD lidar systems continuously record the entire profile of the reflected energy of the emitted pulse over a span of time at pre-selected time intervals (usually 1 to 10 nanoseconds)."*

To understand the practical meaning of the waveform sampling technique (i.e., wave form digitization), I would like to discuss the case of a waveform digitization lidar through a numerical example. A system flown over forested land that has been set to limit the number of samples to 256 samples per pulse, will be able to provide details of a continuous vertical section (depth) equal to:

Scan depth = Number of samples per pulse X vertical resolution of the wave digitizing
= 256 X 0.15m = 38.4 m

extends through the canopy assuming that the fixed sampling interval is set to 1 nanosecond. The laser travels at the speed of light or 299,792,458 meters per second and therefore, the vertical resolution or the distance between consecutive samples is equivalent to the laser ranging in half nanosecond or 0.15 m since the laser travels twice the range. In other words, a layer of information can be collected every 0.15 meters throughout the canopy. To simplify the matter further, when a full waveform digitization capable system is set up to collect a maximum of 256 samples per pulse, it resembles the concept of infinite returns in a multi-return discrete (or conventional) lidar but instead of the few returned pulses (up to five) collected by the discrete lidar, the FWD system samples the returned pulse up to 256 times, assuming that all recorded returns are classified as valid pulse echoes and not noise or falsely detected signals).

Here we need to understand that not all these samples are available for data interpretations and fewer numbers of these samples provide valuable information for the following reasons:

1) If the laser passes through wide gaps in the canopy (not hitting a branch, leaf, trunk, etc.) that is wider than the pulse footprint at that location. In this case, there is no medium to intercept and backscatter the signal resulting in a null data form. However, if there were dense object details (leaves and branches in this case) along the laser path, a stronger backscatter of the signal (or portion of it) is recorded every 0.15 m along 38.4 m of the laser path throughout the canopy for the example given above.
2) Signal noise that degrades the S/N ratio resulting in useless information.
3) Objects that are closely located to each other causing overlapping waveforms resulting in an ambiguous condition to solve accurately. Such condition could result in one of the following effects:
    a) Less accurate ranging for the neighboring objects due to the modified waveforms as illustrated in Figure 1. In Figure 1, although the two objects are physically separated by distance D1, the processed waveforms show them to be apart by distance D2 causing less accurate position determination for both objects.
    b) When an ambiguous condition occurs, the processing software may totally merge the two peaks to show them as one object. Combining wave forms also occurs when the signals are too weak and the software combines them to increase peak detection. This results in less accurate position determination for all combined objects.

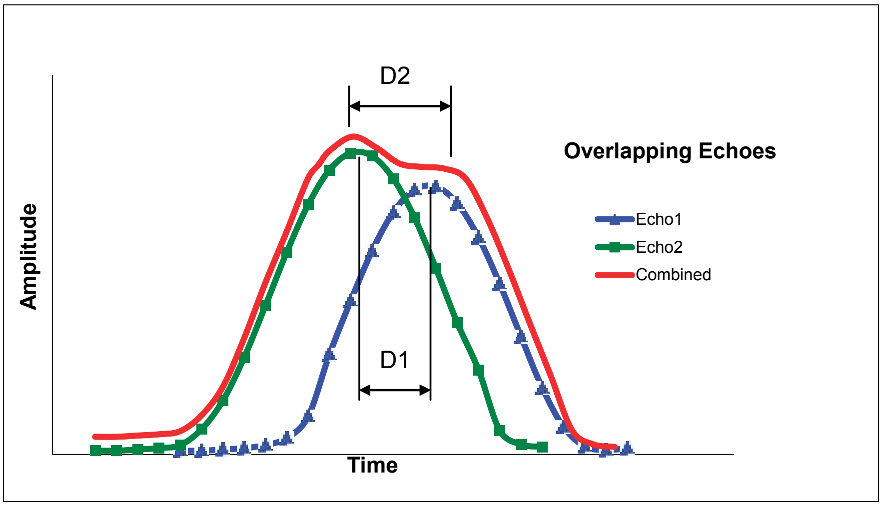

Figure 1. Ambiguity in resolving overlapping echoes.

The level of energy of the returned pulses depends on several factors, which include the following:

1) The physical nature of the object (shape, size, orientation, etc.).
2) How the object physically, and perhaps chemically, interacts with the laser.
3) The strength of the emitted pulse.

Interpretation of waveform data starts with modeling and evaluating the amplitudes of the backscatter at each sample interval. When the laser pulse travels throughout the canopy, for example, amplitudes of the returned echoes are recorded at the pre-set interval during the sampling duration. Larger amplitude is expected when the echo is reflected from a hard object such as a branch or a clump of leaves (Figure 2).

A peak in the returned signal amplitude represents a location on the waveform where strong backscatter is detected (location A of Figure 2). The second peak is expected to show up along the waveform when the laser travels further into the canopy and collides with the next hard object along its path, such as the edge of the leaf in location B of Figure 2. For the peak to be detected at location B, the vertical distance A-B should be larger than the vertical resolution of the system (or 15 cm for the system discussed in my previous example). Amplitudes of the backscatter are recorded for the fragmented pulse through its travel path to the forest floor or until the last sample is recorded, perhaps at location G.

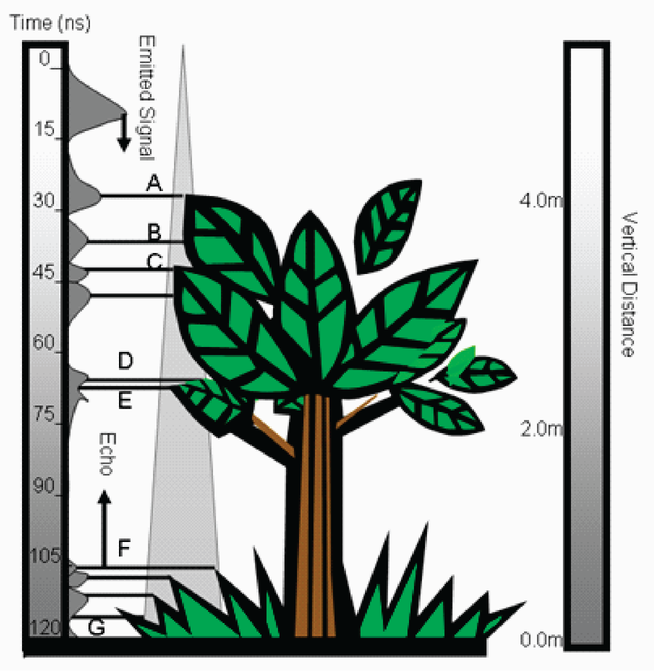

Figure 2. Laser Propagation through Forest.

Analyzing the waveforms provides us with a hint on the shape and the nature of objects along the laser path. In addition, some of the detected peaks can be converted to point cloud resulting in a dense Digital Surface Model (DSM). Objects absorb and/or reflect laser according to the physiological and the physical nature of the object and to its geometry and orientation. All these variations in the backscattered energy or echo become unique signatures for the studied subjects. However, due to some similarity and proximity in the conditions of the canopy and the presence of faint returned signals, in many cases, this will result in overlapping echoes. Overlapping echoes are difficult to detect and separate when using a hardware-based approach that solely relies on thresholding. A better way for observing the backscattered signal is by sampling it with greater frequency and then processing and analyzing the backscattered waveform. The latter technique has proven to be more efficient than thresholding and ultimately results in more detailed information about the studied subject. Comparing data from waveform digitization lidar, if utilized correctly, to data from discrete lidar is similar to comparing cross-sections or profiles through a Digital Topographic Model (DTM) for a levee generated from sub-meter dense lidar dataset to that of the same levee when mapped with 5 meter lidar dataset. Important details of the levee will be lost when using the 5 meter lidar dataset. An efficient signal processing routine to filter out noise from a genuine signal and to maximize the detection of relevant peaks within the signal is crucial to the quality of information derived from FWD lidar. Several techniques were proposed for the task. The most notable and perhaps the simplest of them all is the one based on the simple calculus technique of finding local maxima through the first derivative of the function fitted to the waveform and then verifying its quality through the second derivatives. Besides the range, other valuable characteristics of the backscattered pulse properties are also computed. Those are the pulse (echo) width and the pulse amplitude. Both echo width and amplitude depend on many factors, such as the emitted signal strength, nature and geometry of the illuminated surface, and flying altitude. During the waveform processing, the function $f(x)$, which is fitted to the waveform, is decomposed into components $f_i(x)$ representing the different echoes from objects on the laser path which the signal passes by and echoes from. The decomposed functions should satisfy the relation given in equation 1.

$$f(x) = \sum_{i=0}^{n} f_i(x) \quad \text{.........} \quad (1)$$

A form of Gaussian function is routinely employed by researchers and scientist to represent the summation of the functions $f_i(x)$ in equation 1. Another technique that is often used besides Gaussian decomposition is the cross correlation between the waveforms of the emitted and the returned signals. The later technique (cross correlation) makes distinction between strong and weak returned signals. The weak signal is the one that goes through further processing and re-processing in order to improve its detection. Sometimes several weak returns from two objects that fall in proximity to each other are combined in order to improve the signal quality (Figure 1). When such waveform merging happens, it degrades the accuracy of range determination and therefore results in less accurate location of the two objects. Full

> *“Analyzing the waveforms provides us with a hint of the shape and the nature of objects along the laser path.”*

continued on page 108

continued from page 107

waveform digitization is successfully used in forestry by researchers and scientists for tree species classification. Forest canopy hides many details on the size, species, growth stages, height and condition of the different vegetation that thrive on the forest floor. Waveform analysis is the only remote sensing technique that is proven to be effective in revealing the secrets that lie under the canopy and to map the DNA of the forest floor.

> *"The waveform is in the forefront of the new generation lidar. It made its introduction to the commercial sector and will be slowly replacing the discrete lidar mode collection we are currently using."*

Traditionally, color infrared imagery is used in remote sensing for the classification of tree species, but the method has its limitations in seeing the details buried under the top canopy. lidar data based on full waveform digitization enabled researchers to devise a species classification method that resulted in successful mapping of spruce, beech and larch species, and they were able to separate conifer from broadleaf vegetation.

The last explanation of full waveform digitization should succeed in highlighting the main advantages of the full waveform digitization and that is in providing denser point cloud and substantial information (signature) about the ground surface and what covers it, if the data is analyzed correctly. As you may have realized from the above introduction, the waveform is more than just intensity, as current intensity is based on thresholding technique. Intensity lacks the intricate details that the indefinite digitizing capability of the waveform digitizing system provides about the returned signal or echo from objects along the laser path. While it is easier to analyze point cloud and model objects in their 3-dimentional forms, it is not straight forward, at least for the time being, to interpret the waveform information about ground objects. The field is still a science project in most ways and very few, if any, commercial software packages exist today to enable the user to exploit the full capabilities of the full waveform digitization in remote sensing analysis.

The other issue that is related to the topic and needs to be addressed by the geospatial community in the near future is a standard to define and store the full waveform. For the time being, the LAS 1.4 format is designed to handle the waveform data (See the Highlight Article – The LAS 1.4 Specification – in this issue).

> *"A FWD-enabled system records both discrete returned pulses and additional data to define its waveform."*

It is worth mentioning here that operating a lidar system in FWD mode reduces the maximum pulse rate of some commercial lidar systems. However, considering that waveform data is also used to add additional point cloud to the one collected by the discrete mode, such limitation is marginalized.

Finally, the waveform is in the forefront of the new generation lidar. It made its introduction to the commercial sector and will be slowly replacing the discrete lidar mode collection we are currently using. The mass transition of users to the full waveform digitizing type lidar is a matter of time, and its speed will depend on the development of the tools and software needed to take full advantage of the waveform data. Until such software and tools are fully developed, and end users of lidar data clearly see value in the data collected by a full waveform digitizing lidar, the field will advance sluggishly.

Please send your question to Mapping_Matters@asprs.org and indicate whether you want your name to be blocked from publishing.

Answers for all questions that are not published in *PE&RS* can be found on line at www.asprs.org/Mapping Matters.

***Dr. Abdullah is the Chief Scientist at Fugro EarthData, Inc, Frederick, Maryland. He is the 2010 recipient of the ASPRS Photogrammetric (Fairchild) Award.*

◆

The contents of this column reflect the views of the author, who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the American Society for Photogrammetry and Remote Sensing and/or Fugro EarthData, Inc.