Małgorzata Słota*

Advanced Processing Techniques and Classification of Full-waveform Airborne Laser Scanning Data**

1. Introduction

This paper provides an overview of advanced full-waveform LiDAR data processing techniques. The paper is an extension of the article *Full-waveform Airborne Laser Scanning Data Processing Techniques* [19].

Since full-waveform scanner systems had emerged on the commercial market the potential of full-waveform data has gained a lot of interest.

In general, studies on the potential of full-waveform data can be divided into three groups:

1) Studies focused on the improvement of the classification / segmentation process of point clouds by including additional full-waveform parameters. It is a well known fact that both classification and segmentation rely on an analysis of the differentiability of data sets based on specified attributes values. Typically, these point cloud attributes are related to geometrical relationships between points, e.g. number of points in a defined neighbourhood or height differences between points. Additionally, by modelling raw full-waveform signals, detected echoes can be characterized by shape parameters such as echo width or wave skewness. Advanced modelling aims at increasing the accuracy of echo parameters determination or, alternatively, to appoint new, additional parameters.

2) Analysis aimed at increasing the density of point clouds by the detection of weak and overlapping backscattered laser pulses.

* AGH University of Science and Technology, Faculty of Mining Surveying and Environmental Engineering, Department of Geoinformation, Photogrammetry and Remote Sensing of Environment, Kraków, Poland

3) Research focused on point cloud accuracy improvement. It has been shown that by implementing suitable processing methods of full-waveform data the accuracy of derived point clouds can be significantly increased. Most research in this area focuses on developing improved peak detection methods, which directly corresponds to vertical accuracy enhancement.

Obviously, the categorization of trends in full-waveform data research presented above is purely schematic and only introduces a coarse systematization of studies. In practice, development in one category could be followed by improvement in another category, e.g. an improvement of the signal modelling accuracy of the determined point cloud could potentially increase the density of the point cloud in comparison to classic, discrete LiDAR systems.

In order to maintain a clear article structure, the text has been divided into three main parts. In the first part the advanced signal decomposition techniques along with methods of weak and overlapping pulses detection are presented. In section 3, the idea of radiometric calibration with the most important equations can be found. Since waveform modelling is directly related to full-waveform features extraction for point cloud classification purposes, the subsequent section of the article is dedicated to the classification of LiDAR data. The overview of studies on the possibilities of improving classification results by using point attributes computed from full-waveform energy profiles is given.

2.   Advanced Processing Methods

For most of the full-waveform data, especially for single echoes from planar targets, typical processing methods presented in [19] are sufficient for precise 3D point extraction from the energy profile. However, a critical measurement condition can occur when the reflected pulse is strongly deformed or more than one echo is recorded [8]. The complex signal can be registered for small objects randomly distributed within the laser footprint area, for more than one target with slightly shifted elevation or for an uneven, rough or sloped surface. In such cases, advanced processing is usually required.

The most typical approaches for full-waveform data processing rely on signal decomposition. The idea of decomposition of full-waveform LiDAR data focuses on extracting individual echoes from raw signal, by modelling waves using a basic function. The general formulae of signal decomposition is presented below:

$$y = f(x) = \sum_{j=1}^{n} f_j(x) \tag{1}$$

where:

n – number of components,
f_j – analytical function.

Various sets of parameters can be extracted using different parameter functions in the modelling process. A list of commonly used functions and their characteristics is presented in [19]. The most popular processing method is signal decomposition into Gaussian components. Using Gaussian modelling amplitude and echo width parameters can be retrieved (Fig. 1a, b). Using Generalized Gaussian for signal modelling, an additional flattening parameter is extracted (Fig. 1c). On the other hand, asymmetric functions, such as Lognormal, Weibull or Nakagami, allow us to determine echo skewness (Fig. 1d).
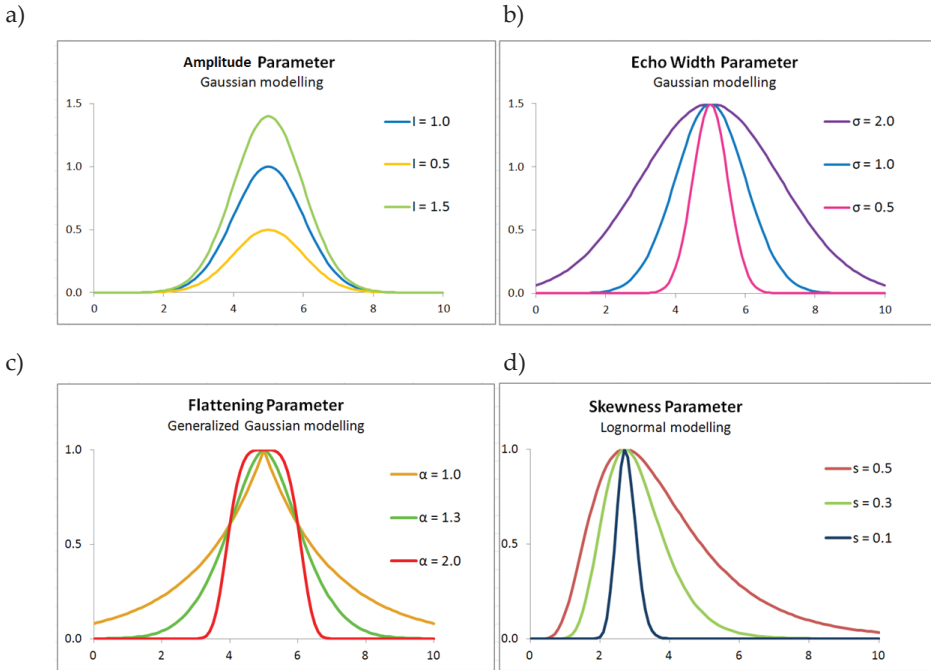

Fig. 1. Shape parameters

As mentioned above, more sophisticated processing methods could be required for data having complex echoes. The overview of some advanced processing techniques is presented below:

- In [25], the authors propose a fast method called PWD (Progressive Waveform Decomposition). This decomposition method does not require initial values and is not iterative. In the first step, the waveform is de-noised and smoothed. Then, a local maximum is detected and a Gaussian function is fitted to the signal. Secondly, the fitted wave is subtracted from the signal and peak detection as well as the Gaussian fitting process are conducted again. The procedure is repeated until no visible peaks are present in the subtracted signal.

- Mallet et al. [13] present a novel signal decomposition approach using the Marked Point Processes. In this method the number of components does not have to be known. The algorithm fits data with functions from a predefined parametric function library.
- The advanced full-waveform decomposition technique for digital elevation model generation purposes was proposed by Hu et al. in [6]. The method starts with a terrain TIN model. For every TIN facet full-waveform profiles intersecting the currently analyzed facet are checked for a peak representing the terrain. The TIN model is successfully updated with newly detected points. This seeded Gaussian decomposition method allows us to detect terrain points under vegetation. The presented preliminary results show that the proposed algorithm can increase the accuracy of the ground model under trees and shrubs by more than 30% [6].
- In [8], authors propose a method of computing a full-waveform signal in the case of different variety of emitted pulses. To remove the influence of the transmitted signal on the recorded backscatter waveforms and calculate the surface response the deconvolution technique is used. The surface response is obtained by transforming the received waveform into the Fourier domain and by applying the Wiener filter for noise reduction.
- Waveform decomposition using a one-dimensional wavelet transform is proposed among others by Molnar et al. [14] and Toth et al. [21]. The biorthogonal CDF 3/9 wavelet is used, which provides good waveform representation and compression.
- The 3D wavelet decomposition technique for vertical object detection and recognition is proposed by Parrish in [15]. After the process of deconvolution and georeferencing of full-waveform data a dense point cloud is obtained. Then, this point cloud is voxelized and a 3D wavelet transform is performed. The objects are detected in the wavelet domain by a multiresolution template matching approach.

By recording entire backscattered energy profiles and post-processing them in an advanced manner additional points could be extracted. However, difficulties arise while computing complex, weak or overlapping echoes, which are likely to occur in vegetated areas [9]. In such places, most of the energy is scattered by the tree canopy layer, resulting in weak echoes for the ground. Most often, such echoes are below the threshold value for typical processing methods. Overlapping echoes, on the other hand, can occur in areas of low vegetation, where the distance between two consecutive targets diminishes. The shapes of superimposed echoes can significantly deviate from single returns, which make correct target extraction very difficult or even impossible for standard pulse detection methods. In fact, typically only one uncertain point between shrubs and terrain is extracted for low vegetation areas, resulting in overestimation of the ground height.

The studies on weak and overlapping echoes were conducted by Stilla et al. [20], Lin et al. [9] and Wang [24].

A summary of the proposed methods is presented below:

– Stilla et al. propose a method of analyzing the local neighbourhood relations by waveform stacking for searching a prominent geometric pattern in waveform data in order to detect weak and overlapping echoes. The algorithm was tested on a regular, simulated data set, which can be expressed as discrete data cuboid I [x y t]. Results have shown that partially occluded objects and targets with poor surface response can be detected when using the proposed method [20].

– The algorithm proposed by Lin is based on the Gaussian decomposition technique. It includes two processing steps. Firstly, Gaussian modelling for visible peaks is applied. Then, shapes of overlapping echoes are analyzed. The algorithm detects inflexion points for asymmetric pulses on both sides of the primary peaks. As a result, overlapping echoes, representing asymmetric behaviour or a non-Gaussian distribution, are successfully resolved. Weak Gaussian pulses are properly extracted by the algorithm [9].

– Wang introduced a detector based on the wavelet transform. The signal is decomposed and possible echoes are extracted from wavelet coefficients at a certain scale. The proposed algorithm identifies all the potential echoes in the registered signal. In the cited study, the influence of noise on the signal as expressed in the signal-to-noise-ratio (SNR) and the ability to detect weak pulses depending on the noise level were examined. For the study area, covering irregularly spaced shrubs and tall crops, more than 18% of additional returns were found using the proposed algorithm [24].

3. Radiometric Calibration

One of the advanced processing techniques related to the decomposition process is the radiometric calibration of the backscattered signal. Calibration of full-waveform data using the radar equation was proposed by Wagner et al. in 2006 [23]. It provides an estimate of the backscatter cross-section for each target and is necessary for a comparison of data acquired by different systems [23] or when analyzing data sets consisting of different flight strips. The calibration equation is presented below:

$$\sigma_i = C_{cal} \cdot R_i^4 \cdot I_i \cdot W_i \tag{2}$$

where:

σ_i – backscatter cross-section for i-th echo,
C_{cal} – calibration constant,
R_i – range (the distance from the sensor to the target),
I_i – echo amplitude,
W_i – echo width.

The calibration constant can be calculated by extracting the amplitude and echo width values for small asphalt areas at nadir view. Under the assumption that these pulses behave like ideal Lambertian scatterers with a reflectivity of $\rho = 0.2$, the estimate of the calibration constant can be expressed as [23]:

$$C_{cal} = \frac{0.2 \cdot \pi \cdot \beta_t^2}{R^2 \cdot I_{asphalt} \cdot W_{asphalt}}$$

(3)

where:

β_t – laser beam divergence,
$I_{asphalt}$ – echo amplitude for asphalt areas,
$W_{asphalt}$ – echo width for asphalt areas.

4. Full-waveform Data for Classification Purposes

The point cloud classification process based on data acquired by the discrete, multi-echo laser system typically operates on geometrical properties (e.g. height difference, co-planarity), or on characteristics such as intensity or the number of returns. Along with the introduction of full-waveform scanner systems, researchers discovered the possibility of extracting additional parameters associated with the shape of the waveform for point cloud classification purposes. Since then, studies on the improvement of classification accuracy by adding parameters directly acquired from full-waveform data into geometrical attributes have been carried out. The main aspect is the existence of a specified set of full-waveform parameters, allowing us to completely replace geometrical attributes in the classification process.

Research in this area has been carried out, inter alia, by Duong et al. [2], Heinzel and Koch [3], Höfle et al. [4, 5], Mallet et al. [10, 11], Molnar et al. [14], Reitberger et al. [17, 18] and Wagner et al. [22]:

– In [2] the possibility of using ICESat data for land-cover classification is described. A decision tree classifier is used to assign the footprints of the laser pulses based on full-waveform parameters into four categories: high vegetation, urban, water, and bare land / low vegetation. A classification accuracy of 73% was achieved.
– The possibility of utilizing full-waveform characteristics for point cloud classification into vegetation and non-vegetation echoes is described in [4, 22]. In [22] the scattering characteristics (width of the backscattered pulse and backscatter cross-section) of vegetation and the underlying terrain are investigated. In [4] a classification based on different sets of segment features computed from the point cloud and derived from full-waveform attributes is proposed. Two machine learning techniques, decision trees and artificial neural networks, were tested in the classification process.

- In [3, 5, 17, 18], the problem of tree species determination based on full-waveform features is investigated. Heinzel and Koch [3] describe an approach to finding the most important factors for tree species classification from full-waveform LiDAR data. Over 200 parameter constellations were explored with linear discriminant analysis (LDA) to determine the most suited variables. Altogether three variables are recommended for the main species classification in Central European temperate forests: the mean intensity, the median of the width and the mean total number of targets within a laser beam. In [5], Höfle et al. propose a new approach for area-based parameterization of the forest structure. The distributions of full-waveform point cloud attributes were analyzed with respect to the backscattering characteristics of individual tree species and forest stand level. In [17, 18] a tree species classification method using salient features is demonstrated. It was shown that classification accuracy for deciduous and coniferous trees can reach 93% and 95% for unsupervised and supervised classification, respectively.
- In [10, 11, 14] the potential of using full-waveform data for classification in urban areas is discussed. In [11] a pattern recognition approach with a Support Vector Machine classification based on full-waveform parameters is proposed. In [10] the relevance assessment of full-waveform features for automatic classification of urban areas is given. The parameters derived from full-waveform data have been ranked using three methods: F-score, Relief and Sum-RFE. Various scenarios based on the rank of features have been tested with a Support Vector Machines classifier. Molnar et al. in [14] have investigated the relevance of parameters computed from waveforms as well. However, different classification methods have been used: supervised Bayes and unsupervised Self-Organizing Map.

As mentioned above, the improvement of point cloud classification can be achieved by adding additional full-waveform features during the classification process. These parameters are associated with the shape of the backscattered laser pulse. On the other hand, the shape of the returned waveform is directly related to the physical properties of the reflecting surface, like surface material or angle of incidence between the scanner and the target [14].

Most of the reflected LiDAR pulses can be sufficiently described by amplitude, echo width, flatness and skewness parameters:

- Amplitude value is related to intensity, registered by traditional laser scanners, and it strongly depends on both the radiometric and geometric properties of the targets. In [11], Mallet et al. show that high amplitudes are observed for building roofs regardless of the material (except metal), on bare earth, and on cars. On the other hand, low values could indicate vegetation (due to a higher target heterogeneity) and street points [11]. This parameter can be useful for the discrimination of buildings from vegetation and ground points.

– Laser pulses travelling through layers of tree branches are spread. That explains why the highest echo width values are observed for vegetation. Ground points and urban waveforms are generally characterised by narrower width values due to the reflectance from sharp or artificial objects [2], even in areas with increasing roof slope [11]. The echo width parameter has proved to be very useful for digital terrain model generation purposes [10]. In [16] the potential of using width for discrimination of ground points from the vegetation points was investigated. It was shown that on average the width of the vegetation points tends to be wider in comparison to the ground points. Nevertheless, a simple classification based only on amplitude and the width value does not have to cause positive results – the low amplitude and the wide echoes do not necessarily indicate vegetation.

The other two shape parameters, flatness α and skewness s, were shown to be rather insignificant for classification purposes [10]. Slightly varying flatness values can be obtained for buildings, ground and vegetation areas. For small-footprint laser data, the waveform asymmetry is irrelevant and is not linked to specified objects [10]. Asymmetric echoes can be noticed for building edges or in areas with surface discontinuities [13].

Other full-waveform parameters found in the literature are: "penetration" and residuals of Gaussian fitting parameters [14], total energy and the beginning of the waveform [2], salient features describing tree structure [17], features from calibration procedures such as backscatter cross-section, the cross-section per illuminated area and the backscatter coefficient [10]. The type of basic function applied at the decomposition stage can be additionally used in the classification procedure, however, it has been shown that a simple symmetric decomposition technique is sufficient for small-footprint data [10].

In summary, the inclusion of full-waveform parameters can improve classification, however, the classification entirely based on waveform shape parameters gives insufficient results [14].

5.   Summary

LiDAR full-waveform data pose a wide range of possibilities for many application purposes, although this enormous potential is still not fully exploited, especially by commercial companies using popular data processing software. This situation will certainly change in the coming years, mainly thanks to many studies conducted in research and academic institutions. One of the main areas of research is the possibility of point cloud classification improvement by including full-waveform parameters. The cited articles indicate that by adding additional full-waveform features related to backscattered waveform shape, such as waveform width, amplitude, flatness or skewness, classification results could be significantly improved. This additional

full-waveform attributes for simple reflections might be extracted by applying typical processing methods like Gaussian signal decomposition. However, proper analysis of complex, weak or overlapping echoes could be a challenging task. A solution is to apply more sophisticated processing techniques, like waveform stacking, wavelet decomposition or the Progressive Waveform Decomposition (PWD) method. While computing data from multiple flight campaigns or from different strips, the radiometric calibration is recommended before the actual data processing.

References

[1] Bretar F., Chauve A., Mallet C., Jutzi B.: *Managing full waveform LIDAR data: A challenging task for the forthcoming years*. The Internal Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVII-B1, 2008, pp. 415–420.

[2] Duong H., Pfeifer N., Lindenbergh R.: *Full waveform analysis: ICESat laser data for land cover classification*. ISPRS Mid-term Symposium, Remote Sensing: From Pixels to Processes, 2006, pp. 8–11.

[3] Heinzel J., Koch B.: *Exploring full-waveform LiDAR parameters for tree species classification*. International Journal of Applied Earth Observation and Geoinformation, vol. 13(1), 2011, pp. 152–160.

[4] Höfle B., Hollaus M., Hagenauer J.: *Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data*. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 67, 2012, pp. 134–147.

[5] Höfle B., Hollaus M., Lehner H., Pfeifer N., Wagner W.: *Area-based parameterization of forest structure using full-waveform airborne laser scanning data*. SilviLaser 2008, 8th International Conference on LiDAR Applications in forest assessment and inventory, Heriot-Watt University, Edinburgh, UK, 17–19 September, 2008.

[6] Hu B., Gumerov D., Wang J.-G.: *An intergation approach to accurate DEM generation using airborne full waveform LiDAR data*. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVIII-5/W12, 2011, pp. 237–241.

[7] Jutzi B., Stilla U.: *Laser pulse analysis for reconstruction and classification of urban objects*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIV-3/W8, 2003, pp. 151–156.

[8] Jutzi B., Stilla U.: *Range determination with waveform recording laser systems using a Wiener Filter*. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 61(2), 2006, pp. 95–107.

[9] Lin Y.-C., Mills J., Smith-Voysey S.: *Detection of weak and overlapping pulses from waveform airborne laser scanning data.* SilviLaser 2008, 8th International Conference on LiDAR Applications in forest assessment and inventory, Heriot-Watt University, Edinburgh, UK, 17–19 September, 2008, pp. 478–487.

[10] Mallet C., Bretar F., Roux M., Soergel U., Heipke C.: *Relevance assessment of full-waveform lidar data for urban area classification*. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 66(6), supplement, 2011, pp. S71–S84.

[11] Mallet C., Bretar F., Soergel U.: *Analysis of full-waveform lidar data for classification of urban areas*. Photogrammetrie Fernerkundung GeoInformation (PFG), vol. 5, 2008, pp. 337–349.

[12] Mallet C., Bretar F.: *Full-waveform topographic lidar: State-of-the-art*. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 64(1), 2009, pp. 1–16.

[13] Mallet C., Lafarge F., Roux M., Soergel U., Bretar F., Heipke C.: *A marked point process for modeling lidar waveforms.* Image Processing, IEEE Transactions on, vol. 19(12), 2010, pp. 3204–3221.

[14] Molnar B., Laky S., Toth C.: *Using Full Waveform Data in Urban Areas.* The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, vol. XXXVIII-3/W22, 2011, pp. 203–208.

[15] Parrish C.E.: *Exploiting full-waveform lidar data and multiresolution wavelet analysis for vertical object detection and recognition*. Geoscience and Remote Sensing Symposium, IGARSS 2007, IEEE International, 2007, pp. 2499–2502.

[16] Persson A., Soderman U., Topel J., Ahlberg S.: *Visualization and analysis of full-waveform airborne laser scanner data.* The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-3/W19, 2005, pp. 103–108.

[17] Reitberger J., Krzystek P., Stilla U.: *Analysis of full waveform lidar data for tree species classification*. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-3, 2006, pp. 228–233.

[18] Reitberger J., Krzystek P., Stilla U.: *Benefit of airborne full waveform lidar for 3D segmentation and classification of single trees*. ASPRS 2009 Annual Conference, Baltimore, Maryland (USA), 2009.

[19] Słota M.: *Full-waveform Airborne Laser Scanning Data Processing Techniques*. Geomatics and Environmental Engineering, vol. 8, no. 1, 2014, pp. 61–74.

[20] Stilla U., Yao W., Jutzi B.: *Detection of weak laser pulses by full waveform stacking*. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-3/W49A, 2007, pp. 25–30.

[21] Toth Ch.K., Zaletnyik P., Laky S., Grejner-Brzezinska D.A.: *Peak detection from full-waveform LiDAR data*. International LiDAR Mapping Forum, New Orleans, Louisiana, 2011.

[22] Wagner W., Hollaus M., Briese C., Ducic V.: *3D vegetation mapping using small-footprint full-waveform airborne laser scanners*. International Journal of Remote Sensing, vol. 29(5), 2008, pp. 1433–1452.

[23] Wagner W., Ullrich A., Ducic V., Melzer T., Studnicka N.: *Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner*. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 60(2), 2006, pp. 100–112.

[24] Wang C.K.: *Exploring weak and overlapped returns of a lidar waveform with a wavelet-based echo detector*. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIX-B7, 2012, pp. 529–534.

[25] Zhu J., Zhanga Z., Hu X., Lia Z.: *Analysis and application of LiDAR waveform data using a progressive waveform decomposition method*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVIII-5/W12, 2011, pp. 31–36.