7-2007

# Exploiting Full-Waveform Lidar Data and Multiresolution Wavelet Analysis for Vertical Object Detection and Recognition

Christopher Parrish
*University of New Hampshire, Durham*

# Exploiting Full-Waveform Lidar Data and Multiresolution Wavelet Analysis for Vertical Object Detection and Recognition

Christopher E. Parrish
NOAA, National Geodetic Survey
1315 East-West Hwy, Silver Spring, MD 20910

***Abstract*—A current challenge in performing airport obstruction surveys using airborne lidar is lack of reliable, automated methods for extracting and attributing vertical objects from the lidar data. This paper presents a new approach to solving this problem, taking advantage of the additional data provided by full-waveform systems. The procedure entails first deconvolving and georeferencing the lidar waveform data to create dense, detailed point clouds in which the vertical structure of objects, such as trees, towers, and buildings, is well characterized. The point clouds are then voxelized to produce high-resolution volumes of lidar intensity values, and a 3D wavelet decomposition is computed. Vertical object detection and recognition is performed in the wavelet domain using a multiresolution template matching approach. The method was tested using lidar waveform data and ground truth collected for project areas in Madison, Wisconsin. Preliminary results demonstrate the potential of the approach.**

***Keywords-lidar; waveform; wavelet; 3D; object detection; airport obstruction surveys***

## I. INTRODUCTION

The U.S. National Geodetic Survey (NGS) manages an aeronautical survey program, which provides airport obstruction survey data to the Federal Aviation Administration (FAA). The primary goal in an airport obstruction survey is to accurately geolocate vertical objects, such as trees, poles, antennas, towers, and buildings, that penetrate obstruction identification surfaces (OIS) enveloping the airfield and approach paths. NGS airport obstruction surveys are currently performed through a combination of photogrammetric and field survey techniques, which have been proven to yield very accurate, reliable obstruction data [1]. However, these methods are also time-consuming and costly. These factors have precipitated an increasing interest, particularly within the private sector, in using emerging remote sensing technologies to increase efficiency and reduce costs.

One remote sensing technology that appears especially well suited for this application is airborne lidar (also referred to as "airborne laser scanning"), which enables accurate, efficient mapping of terrain and elevated features, such as tree canopy. However, while previous studies have demonstrated that lidar can be used to meet certain airport obstructions surveying standards [1-3], some problems remain unsolved. Most significantly, reliable, automated methods for extracting and attributing airport obstructions from lidar data do not currently exist. Therefore, at present, lidar airport obstruction surveys require extensive manual labor, negating some of the intended benefits [3].

The work presented here aims to solve this problem through the application of novel processing and analysis techniques that take advantage of recent advances in lidar technology. Unlike earlier commercial lidar systems, which only recorded a small number of returns (e.g., one to four) per transmitted laser pulse, many of the latest commercial systems record the entire return waveform at very high sampling frequencies (~1 GHz), typically yielding a few hundred digitized samples per pulse [4]. The methodology described here was specifically designed to exploit both the additional data captured by these full-waveform systems and the benefits of 3D wavelet analysis for vertical object detection and recognition.

## II. METHODS

The approach to automatically extracting and attributing vertical objects, such as airport obstructions, from full-waveform lidar data is illustrated graphically in Fig. 1. It can be summarized as follows: 1) the lidar waveform data are deconvolved and georeferenced to produce dense, detailed $X,Y,Z,I$ point clouds (where $X$, $Y$, and $Z$ denote spatial coordinates in the mapping frame and $I$ represents intensity); 2) the output point clouds are voxelized to generate high-resolution volumes of lidar intensity values; 3) 3D wavelet decompositions of the volumes of lidar intensity values are computed; and 4) vertical object detection and recognition is performed in the wavelet domain using a multiresolution template matching approach.

The motivation for deconvolving the waveform data is that the desired (ideal) signal consists of a train of spikes in time, where each spike corresponds to an individual laser reflection, with its amplitude proportional to the amount of backscattered energy. For example, in a forested area, the first spike in a particular idealized return might correspond to the reflection from the topmost branch or leaf of a tree, with the amplitude of the spike being proportional to the strength of the return from that surface. Likewise, the second spike might correspond to a reflection from the next lower branch, and so forth. An actual received waveform can be considered a blurry, noisy version of this ideal signal, as expressed in the observation model

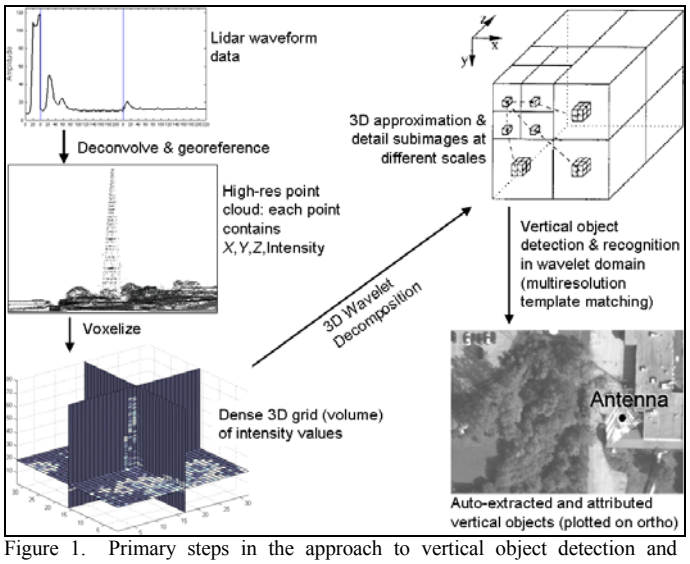


Figure 1. Primary steps in the approach to vertical object detection and recognition in full-waveform lidar data.

$$y[n] = x[n] * g[n] + \eta[n]. \quad (1)$$

In (1), $y[n]$ is the observed waveform, $x[n]$ is the ideal signal, $g[n]$ is a degradation function, $\eta[n]$ is additive noise, and "$*$" denotes convolution.

In this study, two different signal restoration approaches were investigated for "recovering" the ideal signal from the observed waveform: Wiener filtering and the expectation-maximization (EM) deconvolution algorithm described in [5]. The Wiener filter was determined to have serious drawbacks for this application. Specifically, the signals are not well characterized in the frequency domain, leading to significant difficulty in achieving the desired level of deblurring and denoising of the observed waveforms using the Wiener filter approach. Additionally, the Wiener-filtered waveforms exhibited pronounced ringing (artificial oscillations around sharp edges), resulting in false points above and below the ground and objects in the output lidar point clouds.

The EM deconvolution algorithm of [5] enables these limitations to be overcome by utilizing the additional information that the ideal signal (modeled as a random set of spikes in time) is sparse. This algorithm numerically calculates the maximum *a posteriori* (MAP) estimate of the signal by alternating between Fourier-based estimation (the E-step) and denoising in the wavelet domain (the M-step). Our implementation parallels the formulation in [6], except that the denoising M-step is actually performed directly in the time domain, since the time domain representation of the ideal signal (spike train) is already as sparse as possible. A favorable feature of this algorithm is the provision of a tunable parameter, $\tau$, that directly controls the number of spikes in the output signal and, hence, the density of the resulting point cloud. Increasing $\tau$ decreases the point density but also reduces the number of false alarms. Decreasing $\tau$ has the opposite effect. For applications such as airport obstruction surveying in which the consequences of a miss (i.e., an obstruction that the software fails to detect) are more severe than the consequences of a false alarm, $\tau$ can be decreased to produce a more conservative detector [7].

Fig. 2 shows an example of an observed waveform deconvolved using both the Wiener filter and EM algorithm. As illustrated in this figure, the EM algorithm achieves much better results than the Wiener filter; the deconvolved signals are much sharper, the noise has been suppressed to a much greater extent, and the artifacts produced by the Wiener filter (i.e., ringing) do not occur. After deconvolution, the output spike trains are georeferenced using the laser geolocation equation of [2] to generate dense, detailed point clouds in *X,Y,Z,I* format.

The justification for volume representation (i.e., the 3D grid of lidar intensity values depicted in Fig. 1) is twofold. First, it avoids the drawbacks of surface representations, such as digital surface models (DSMs) and triangulated irregular networks (TINs), which are the most commonly-used data models in lidar processing and analysis. These types of surface representations cannot adequately represent vertical structure, because only one elevation per (*X,Y*) coordinate pair is permitted (discontinuities in *Z* and concave vertical shapes are not allowed) [8]. This leads to loss of information about the interior of tree canopies, building sidewalls, roof overhangs, and other types of vertical structure. Furthermore, when using these surface models, it can be difficult to distinguish between erroneous elevations (e.g., due to a bird or electronic noise) and an actual object, such as a pole. The second advantage of the volume representation is facilitation of 3D wavelet analysis. Potential drawbacks of the volume representation, such as loss of information due to interpolation and large file sizes, are mitigated through 3D wavelet representation.

2D wavelet-based approaches to segmenting lidar data into ground ("bare earth") and objects are discussed in [9-12] and related works by these authors. A unique characteristic of this new approach is utilization of a 3D wavelet transform that has been custom designed to facilitate extraction of objects with certain horizontal and vertical structures. Although the design and details of the scaling and wavelet filters used in this research are beyond the scope of this overview paper, three important properties of our wavelet representation can be summarized as follows: 1) very good localization in space, achieved through use of very short (3-tap) filters; 2) detail coefficients that record sharp transitions in intensity, but disregard linear transitions and constant (flat) areas; and 3) simple, fast inversion. Fig. 3 shows a two-scale wavelet decomposition tree based on this 3D wavelet representation.

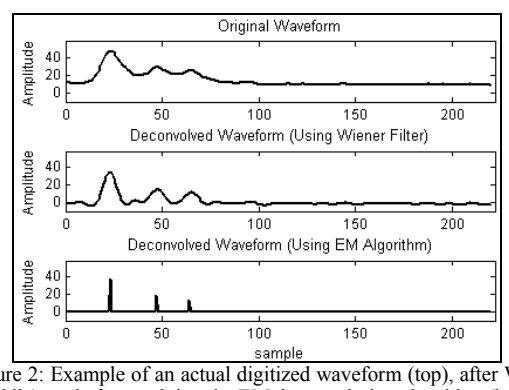


Figure 2: Example of an actual digitized waveform (top), after Wiener filtering (middle), and after applying the EM deconvolution algorithm (bottom).

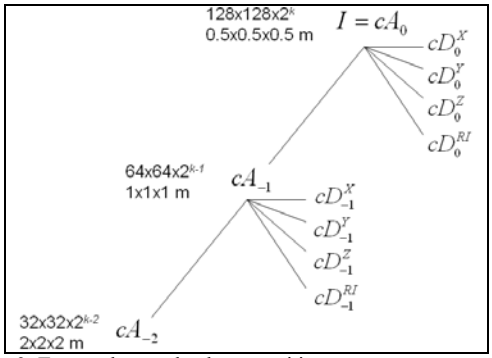

Figure 3: Two-scale wavelet decomposition tree.

In Fig. 3, $cA_j$ denotes level-*j* approximation coefficients; $cD_j^X$, $cD_j^Y$, $cD_j^Z$, and $cD_j^{RI}$ represent horizontal (*X* and *Y*), vertical (*Z*) and rotationally-invariant (*RI*) detail coefficients at level *j*, respectively; and *I* is the original 3D lidar intensity image, which is used as the highest-scale (level 0) approximation. In the vertical object detection/recognition algorithm described below, a two-scale decomposition is computed, with the typical volume sizes and resolutions listed in Fig. 3.

Fig. 4 shows an example of the first level of wavelet decomposition, where the input consists of the volume of lidar intensity values for a tower in one of the Madison project areas. The original 3D lidar intensity image (top left in the figure) serves as the approximation at the highest scale (level 0). The approximation subimage at level -1 is shown in the top right, while the highest-scale composite detail image is shown in the bottom left. The level-*j* composite detail coefficients, $cD_j$, are defined as

$$cD_j = cD_j^X + cD_j^Y + cD_j^Z + cD_j^{RI} \tag{2}$$

The final step in the approach is multiresolution template matching by normalized cross-correlation. The basic idea behind template matching is quite straightforward: a reference template of the object to be detected is translated pixel-by-pixel (or voxel-by-voxel in 3D) through the intensity image, and, at each location, a similarity or difference measure is computed. One common similarity measure is the so-called normalized cross-correlation (NCC) defined as

$$c_N(i,j,k) = \frac{\sum_{x,y,z}\left(f(x,y,z) - \bar{f}_{i,j,k}\right)\left(t(x-i,y-j,z-k) - \bar{t}\right)}{\left[\sum_{x,y,z}\left(f(x,y,z) - \bar{f}_{i,j,k}\right)^2 \sum_{x,y,z}\left(t(x-i,y-j,z-k) - \bar{t}\right)^2\right]^{1/2}} \tag{3}$$

In (3), *f* is the 3D intensity image, *t* is the reference template, $\bar{t}$ is the mean of *t*, and $\bar{f}$ is the mean of *f* in the sub-volume overlapped by the template. The denominator in (3) reduces the sensitivity to within-scene variation in signal energy (due, for example, to variation in flying height) and normalizes the output such that $c_N$ is always in the interval $[-1,1]$.

Although NCC-based template matching is typically implemented in the spatial domain, it can also be performed directly in the wavelet domain (e.g., on a wavelet subimage) [13, 14]. In this work, a wavelet domain implementation is used, wherein the correlation is computed directly between the magnitudes of the wavelet coefficients of the 3D lidar intensity image and those of the template. That is, *f* and *t* in (3) are the absolute values of the composite detail coefficients (2) for the 3D lidar intensity image and reference template, respectively. Benefits of the wavelet domain implementation include increased computational efficiency using a coarse-to-fine matching strategy and improved performance, since the wavelet representation has been specifically designed to facilitate identification of airport obstructions. At the coarse resolution (level -1), a relatively crude, conservative detection strategy is implemented by using generic (artificial) templates consisting of vertical cylinders of constant value 1. At locations where a detection is registered, matching is repeated at level 0 (the original resolution of the 3D lidar intensity image) using templates for various vertical objects (e.g., deciduous and coniferous trees, poles, towers, and buildings) created from training data and stored in a library. To account for differences in spatial orientation, each template was rotated about the *Z*-axis in 20° increments, and each rotated template was added to the library.

## III. EXPERIMENT

To test the described approach, lidar waveform data were collected over two project areas in Madison, Wisconsin on June 23, 2006 using an Optech ALTM 3100 and waveform digitizer. The data were acquired at a flying height of 800 m (AGL), with an average flying speed of 70 m/s, a pulse repetition frequency of 70 kHz, a scan angle of ± 17.3° and a scan frequency of 49.8 Hz. One of the two project areas was used for training (e.g., designing the template library used in the detection/recognition algorithm), while the other was used for testing. In addition to lidar waveform data, high-resolution aerial imagery was acquired for the same two project areas using a medium-format, directly-georeferenced digital camera onboard a NOAA aircraft. The imagery, post-processed GPS/IMU data, and a 1/3-arcsecond USGS DEM were used to produce 0.3-m resolution orthophotos. Lastly, field surveys were performed to obtain ground truth for 35 vertical objects in each of the two project areas.

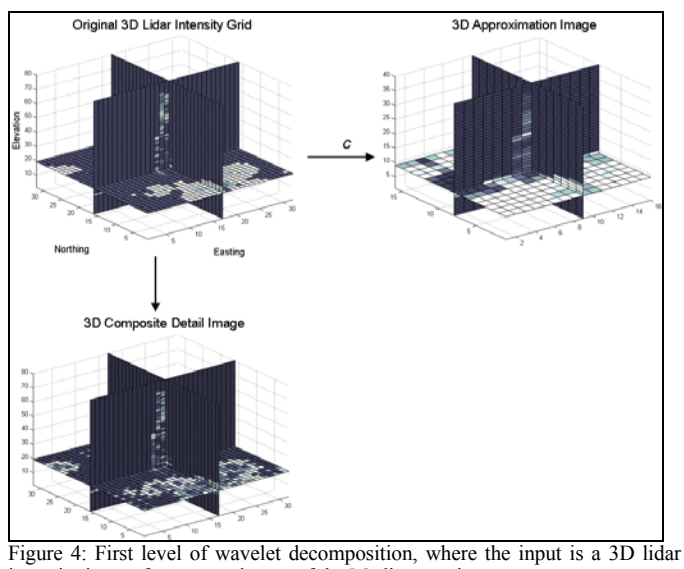

Figure 4: First level of wavelet decomposition, where the input is a 3D lidar intensity image for a tower in one of the Madison project areas.

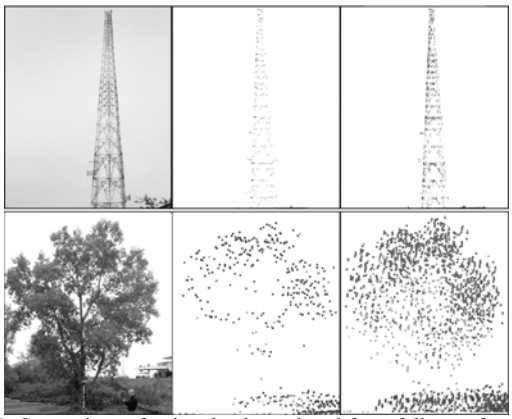

Figure 5: Comparison of point clouds produced from full-waveform lidar data using the algorithms described here (right column) with point clouds generated from discrete-return data processed using the manufacturer's software (middle column). The objects are a tower and tree, as shown in the photos at left.

## IV. Results and Analsyis

To test whether full-waveform lidar data can be used to increase the amount of useful information for airport obstruction surveying, point clouds generated through the first step in Fig. 1 were compared with point clouds generated from discrete-return data output from the manufacturer's software. The results showed that the full-waveform data, processed using the described algorithms, yielded a 252% increase in the average number of points on vertical objects in the test site, with a negligible increase in the amount of noise. Examples are shown in Fig. 5.

The final step is to investigate whether the volume representation and 3D wavelet analysis (steps 2-4 of Fig. 1) lead to better vertical object detection and recognition results than conventional approaches using surface representations. To this end, the results are being compared against those obtained using a DSM generated from the Madison lidar data and the vertical object detection software described in [15]. Preliminary results indicate that the described methods increase detection rates by over 15%. Classification accuracy is still being assessed, as the template library used in the detection and recognition software is not yet complete. The software described in [15] differs from ours in several key ways: namely, it is efficient, user-friendly commercial software, and it has been optimized for InSAR-derived DSMs, rather than raw lidar point clouds. However, the large increase in detection rates indicates that, for our particular application of lidar airport obstruction surveying, volume representations and 3D wavelet analysis can enhance obstruction detection.

## V. Summary and Conclusions

A new methodology for detection and recognition of vertical objects in lidar data has been described. The approach was specifically designed to take advantage of both the additional information contained in full-waveform data and the benefits of multiresolution wavelet analysis. Preliminary results indicate that the approach has significant potential for airport obstruction surveying applications. Future work might involve testing in a wider variety of project areas and operational conditions. Additionally, the approach may be adapted for detection of bottom hazards in bathymetric lidar data.

## References


[1] C.E. Parrish, G.H. Tuell, W.E. Carter, and R.L. Shrestha, "Configuring an Airborne Laser Scanner for Detecting Airport Obstructions," *Photogrammetric Engineering and Remote Sensing*, vol. 71, no. 1, 2005.

[2] C.E. Parrish, "Analysis of Airborne Laser-Scanning System Configurations for Detecting Airport Obstructions," Masters Thesis, University of Florida, Gainesville, Florida, 108 p., 2003.

[3] C. Parrish, J. Woolard, B. Kearse, and N. Case, "Airborne LIDAR Technology for Airspace Obstruction Mapping," *Earth Observation Magazine (EOM)*, vol. 13, no. 4, 2004.

[4] R. Gutierrez, A. Neuenschwander, and M.M. Crawford, "Development of Laser Waveform Digitization for Airborne LIDAR Topographic Mapping Instrumentation," *Proceedings of the IEEE IGARSS Conference*, vol. 2, pp. 1154-1157, July 2005.

[5] M.A.T. Figueiredo, and R.D. Nowak, "An EM Algorithm for Wavelet-Based Image Restoration," *Proceedings of the IEEE Transactions on Image Processing*, vol. 12, no. 8, 2003.

[6] A. Jalobeanu, R.D. Nowak, J. Zerubia, and M.A.T. Figueiredo,. "Satellite and Aerial Image Deconvolution Using an EM Method with Complex Wavelets." *Proceedings of the International Conference on Image Processing (ICIP)*, September, 2002.

[7] C.E. Parrish, and F.L. Scarpace, "Investigating full-waveform lidar data for detection and recognition of vertical objects," American Society of Photogrammetry and Remote Sensing (ASPRS) Annual Conference, May 2007, in press.

[8] J. Stoker, "Voxels as a Representation of Multiple-Return Lidar Data," *Proceedings of the American Society of Photogrammetry and Remote Sensing (ASPRS) Annual Conference*, May 2004.

[9] H. Wei, and M. Bartels, "Unsupervised Segmentation using Gabor Wavelets and Statistical Features in LIDAR Data Analysis," *Proceedings of the 18th International Conference on Pattern Recognition*, vol. 1, pp. 667-670, August 2006.

[10] T.T. Vu, R. Yokoyama, F. Yamazaki, and M. Tokunaga, "Wavelet-based system for classification of airborne laser scanner data," *Proceedings of the IEEE IGARSS Conference*, vol. 7, pp. 4404-4406, July 2003.

[11] M.J. Falkowski, A.M.S. Smith, A.T. Hudak, P.E. Gessler, L.A. Vierling, and N.L. Crookston, "Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data," *Canadian Journal of Remote Sensing*, vol. 32, no. 2, pp. 153-161, 2006.

[12] T. Amgaa, "Wavelet-based analysis for object separation from laser altimetry data," Masters Thesis, International Institute for Geo-Information Science and Earth Observation, Enschede, the Netherlands, 48 p., 2003.

[13] S. Kandadai, and C.D. Creusere, "An experimental study of object detection in the wavelet domain," *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1620-1623, November 2003.

[14] J. Thornton, P. Hennings, J. Kovacevic, and B.V.K. Vijaya Kumar,. Wavelet packet correlation methods in biometrics. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, March 2005.

[15] D. Woods, C. Folley, Y.-T. Kwan, and B. Houshmand, "Automatic Extraction of Vertical Obstruction Information from Interferometric SAR Elevation Data," *Proceedings of the IEEE IGARSS Conference*, vol. 6, pp. 3938 - 3941, September 2004.