

# Fourier Error Analysis of Ray Tracing on a Geospatial Polygonal Model

Brandon J. Baker, *Student Member, IEEE*

**Abstract**— This paper demonstrates that the error associated with ray tracing photorealistic polygonal models of Light Detection and Ranging (LIDAR) scan data is negligible for many applications. In some cases, the standard deviation of error is actually reduced by ray tracing to points on a polygonal model between LIDAR data rather than capturing more data from a scanner. Numerical analyses on the sources of error were performed using Fourier analysis, Taylor expansion, and a statistical model of LIDAR data. Ray tracing was then used to calculate the intersection of each polygon at a point of interest, allowing 3D data to be quantified at any location in between LIDAR points. Actual acquired from a LIDAR scanner and modeled data are then compared to the ray traced values. Results show that many data types have sufficiently low spectral content, allowing accurate representation of 3D data acquired from a LIDAR scanner as a polygonal model.

**Index Terms**—Light Detection and Ranging (LIDAR), error analysis, sensor signal processing, digital terrain model (dtm)

## I. INTRODUCTION

THERE is a need for greater understanding of the error associated with the processing of data acquired by light detection and ranging (LIDAR) scanners [1], even though LIDAR scanners have acquired 3D spatial data for surveys, planning, development, inventory control and other applications for over a decade. An error analysis of modeling LIDAR data from terrestrial scanners is presented in this paper.

When a particular application requires additional accuracy, the solution typically involves scanning at a higher resolution of 3D spatial data. This reliance on data directly acquired from LIDAR scanners has forced many users to gather redundant data at the expense of efficiency. This paper explores the sources of error associated with LIDAR data and the effects of ray tracing polygonal LIDAR data as an alternate to acquiring and utilizing high density scans. The errors are quantified for a paved surface; however the results may be applied to many other types of surfaces.

The results of this analysis demonstrate that polygonal

models of terrestrial LIDAR data are accurate methods for representing DTMs. The results further demonstrate that the accuracy of a polygon created from LIDAR data can be higher than scanning at a higher resolution for many applications.

## II. PREVIOUS RESEARCH

Many researchers have performed error analyses on various algorithms that process aerial LIDAR scanning data [1]-[8]. In reference [1], the authors analyze various algorithms used to estimate spatial data from LIDAR data to create DEMs and DTMs. These previous works describe observations of error in specific environments but do not provide a complete mathematical basis to address the reasons why the particular results were observed. Additional publications discuss the processing of airborne LIDAR data for DTM extraction [9]-[11]. Reference [9] analyzes LIDAR data DTM extraction via adaptive processing; [10] discusses modeling of LIDAR waveforms in vegetation populated terrains; [11] offers insight into techniques for analyzing bathymetric LIDAR data.

This paper explains the mathematics behind estimation theory as it applies to LIDAR scan data, and offers several illustrative examples of the errors associated with actual and modeled scan data to make an accurate assessment of the strengths and limitations of ray tracing on polygonal models.

## III. DATA DESCRIPTION AND ANALYSIS

### A. Initial Error Estimates Using Actual LIDAR Data

The LIDAR data chosen for this project are taken from a paved surface. Two separate LIDAR scans taken from the same location were captured a few minutes apart. Data from one scan were thinned and polygonal models were generated to represent a surface. The discrepancy between the acquired LIDAR data and the ray trace calculation were determined. The difference between LIDAR points from one scan to another, captured from the same location, was also computed. The differences were plotted as a function of distance from the scanner, up to 25m, as shown in Fig. 1. The LIDAR to LIDAR difference exceeded that of the ray trace to LIDAR difference for every distance plotted.

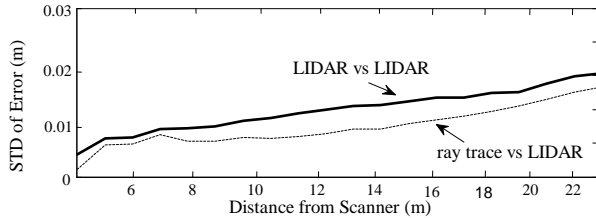


Figure 1. Standard deviation of error vs. distance for the LIDAR vs. LIDAR and ray trace vs. LIDAR examples

The discrepancies between the two plots found in Fig. 1 are the result of low pass filtering. Inherent noise in LIDAR scanners can be reduced on relatively flat surfaces by creating polygons from lower resolution LIDAR points.

#### IV. ANALYSIS OF THE SOURCES OF ERROR

##### A. Overview

In general, there are fundamentally three sources for error when acquiring data. The first is discretization; this source was quantified using Fourier analysis. Second, the method applied in utilizing the data (connecting the dots), which was quantified using Taylor’s expansion. Third is the inaccuracy inherent in the acquisition process, which was analyzed using statistical methods. Models of anomalous elements, a groove, a bump, and an inclined plane, were created and the corresponding Fourier spectra were analyzed.

##### B. Fourier Analysis of Error Due to Finite Sampling Rate

Using one dimensional Fourier analysis, a finite number of harmonics ( $N$ ) can construct a function within an error of  $\epsilon_N$ ,

$$\epsilon_N = \sum_{n=N+1}^{\infty} A_n \cos(\omega_n \phi) + B_n \sin(\omega_n \phi) \quad (1)$$

$\omega_n$  is the spatial frequency,  $\phi$  is the azimuth angle, and  $A_n$  and  $B_n$  are the amplitudes of the even and odd functions in the series, respectively. Dimensions of the various models are given in Table I. Fig. 2 and Fig. 3 depict a cross section of the groove with its corresponding Fourier spectra, respectively. The standard deviation of various errors due to band-limiting, (1), of modeled data for the anomalous groove were calculated for various sampling rates ( $F_s$ ) and distances ( $d$ ), and are given in Table II below.

TABLE I.

| Anomaly | Width  | Height |
|---------|--------|--------|
| Groove  | 428mm  | 21mm   |
| Bump    | 8.56mm | 11.5mm |
| Incline | 500mm  | 25mm   |

TABLE II.

| $F_s$       | $d$ | $\epsilon_N$              |
|-------------|-----|---------------------------|
| 1,000 S/rad | 2m  | $6.268 \times 10^{-17}$ m |
| 1,000 S/rad | 5m  | $2.0163 \times 10^{-6}$ m |
| 1,000 S/rad | 10m | $6.2542 \times 10^{-4}$ m |
| 1,000 S/rad | 50m | $2.86 \times 10^{-2}$ m   |

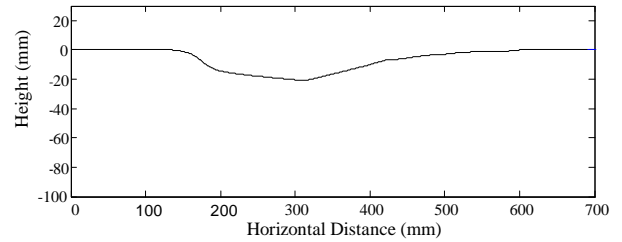


Figure 2. Model of an anomalous groove in a paved surface

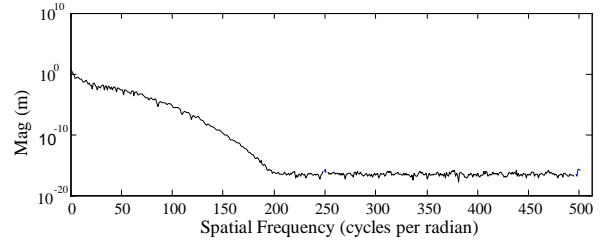


Figure 3. Fourier spectrum of anomalous groove from Figure 2

At 50m the standard deviation of the band-limiting error exceeds that of the maximum height of the anomaly itself. However, at that same sampling rate, the standard deviation of band-limiting error is less than a millimeter when capturing at or below a distance of 10m.

The bump modeled for this project is similar to the rough surface of pavement due to the shape of the pieces of gravel that are compacted as the road is made. The calculated errors of ray tracing polygonal approximations of the bump are given in Table III. These errors suggest that one must scan at an extremely high resolution from two meters away to achieve sub-millimeter accuracy when modeling a small anomaly.

The slightly curved incline in the paved surface is used to represent the slight crowning of a roadway that elevates towards the center of the road. The incline rises approximately 15mm over a span of 1m. The computed errors for the incline are given in Table IV. The errors in Table IV seem extraordinarily high. However, this simply suggests that the reconstruction of an inclined surface should be accomplished by a means other than Fourier. The next section shows that the linear approximation is much better than the band-limited approximation for the modeled incline.

TABLE III.

| $F_s$        | $d$ | $\epsilon_N$             |
|--------------|-----|--------------------------|
| 10,000 S/rad | 2m  | $3.278 \times 10^{-4}$ m |
| 10,000 S/rad | 3m  | $1.153 \times 10^{-3}$ m |
| 10,000 S/rad | 5m  | $2.717 \times 10^{-1}$ m |
| 10,000 S/rad | 10m | $5.00 \times 10^{-1}$ m  |

TABLE IV.

| $F_s$       | $d$ | $\epsilon_N$              |
|-------------|-----|---------------------------|
| 1,000 S/rad | 2m  | $6.268 \times 10^{-17}$ m |
| 1,000 S/rad | 5m  | $2.0163 \times 10^{-6}$ m |
| 1,000 S/rad | 10m | $6.2542 \times 10^{-4}$ m |
| 1,000 S/rad | 50m | $2.86 \times 10^{-2}$ m   |

### C. Error Due to Linearization

The preceding error estimation in (1) assumes that the function,  $f(\phi)$  is approximated using sine's and cosines. However, for this project, the function  $f(\phi)$  is represented by planar surfaces (triangulated 3D polygons), which can be reduced to lines when analyzing LIDAR data as a function of azimuth angle only.

The error of a linear approximation to any function can be quantified by using Taylor's expansion of the Fourier sine and cosine series. Taking the nonlinear terms of a particular frequency of cosine or sine, respectively, yields for each  $A_n$  and  $B_n$ , respectively, the total error associated with the polygonal (linear) approximation. Thus,

$$\epsilon_{Lin} = \sum_{n=1}^N A_n \sum_{m=1}^{\infty} \frac{(\omega_n \phi)^{2m}}{(2m)!} + \sum_{n=1}^N B_n \sum_{p=1}^{\infty} \frac{(\omega_n \phi)^{2p+1}}{(2p+1)!} \quad (2)$$

the sum of which constitutes the linearization error for odd and even, non-linear functions, respectively.

The total error, then, of the polygonal model versus the true paved surface can be quantified by summing the errors from (1) and (2). Finally, the surface of the paved surface,  $f(\phi)$ , can be represented as a function of the terms used in the polygonal (linear) approximation and the total error, as

$$f(\phi) = \sum_{n=1}^N (A_n + B_n \omega_n \phi) + \epsilon_N + \epsilon_{Lin} \quad (3)$$

A few of the results are given in Table V. The standard deviation of the total error for each model shows that the ill-effects of polygonalization can be neglected in many cases. In cases where the total error is less than the error due to band-limiting, one can assume that the linear approximation to the surface is a better approximation than the Fourier reconstruction, and thus the error is reduced when linearized.

The standard deviation of the total error for the anomalous bump as captured from  $2m$  away with a sampling frequency of 1,000 S/rad was calculated to be  $0.599mm$  (as shown in Table V). The linear approximation and original model for the anomalous bump is shown below in Fig. 4.

One can recognize from Fig. 4 where the largest errors are due to the polygonalization of LIDAR data. However, the standard deviation for these particular settings is still less than a millimeter. The results of this numerical analysis suggest that polygonalization is not the primary source of error in scan data. Rather, limitations due to noise in the scanning hardware contribute far more to the overall error in the data than polygonalization. In fact, polygonalization can actually reduce the error of scan data for certain objects.

TABLE V.

| Anomaly | $F_s$       | $d$ | $\epsilon_N + \epsilon_{LIN}$ |
|---------|-------------|-----|-------------------------------|
| Groove  | 1,000 S/rad | 50m | 0.474mm                       |
| Bump    | 1,000 S/rad | 2m  | 0.599mm                       |

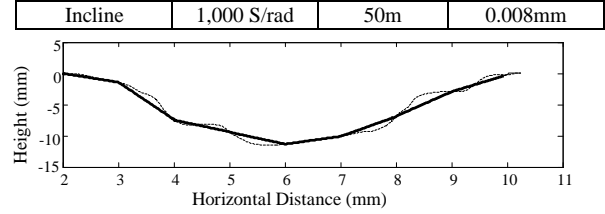


Figure 4. Modeled surface anomaly (bump) and linear approximation

### D. Error Due to Scanner Noise

The third source of error is the noise inherent in the scanning hardware. The polygonal model of the inclined plane shown in Fig 5 had a standard deviation of error of  $10mm$ . The noise in the data can be easily seen in the Fourier spectrum of a sample of data, as shown in Fig. 6.

By comparing the results of the data in Fig. 6 to the model in Fig. 5 of paved surface data, it can be readily seen that there is noise in the higher frequency bands in the actual data. If a perfectly flat surface were scanned, a time of flight scanner would still have errors from  $6mm$  to  $15mm$ , without further processing.

Contrastingly, reducing the sampling rate of the LIDAR data by polygonalization effectively removes the high frequency noise for relatively flat surfaces. Anomalous objects as small as or smaller than the noise region are difficult to detect in a single scan due to the random distribution of frequency components present in the noise. However, utilizing the statistical law of large numbers one could uncover anomalous elements within a series of scans of the same object or scene, even if the size of the object(s) were smaller than the level of noise.

### E. The Law of Large Numbers

The Law of Large Numbers mentioned above describes the evolution of statistical behavior as the number of samples in a set of random variables increases. It states that the average of a set of random variables approaches its statistical mean as the number of samples in the set increases.

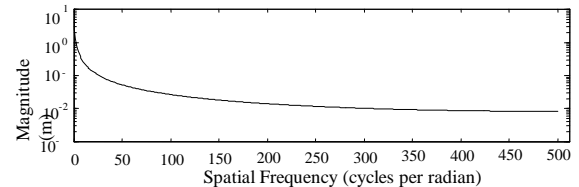


Figure 5. Fourier spectrum of the gradual incline

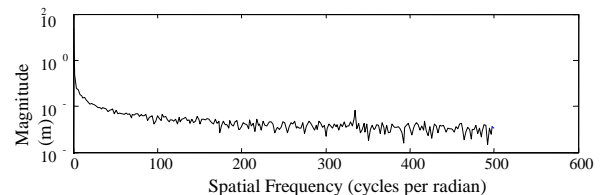


Figure 6. Fourier spectrum of a single horizontal slice of LIDAR data on

a paved surface

To illustrate this concept, a set of samples of normally distributed random data, representing range data captured by a scanner with a standard deviation of error of 6mm, was analyzed. The errors of the first sample were not averaged. The errors of the next three sets of samples were then analyzed taking averages of 10, 100, and 1,000 points, respectively. A normalized probabilistic curve for each sample set is shown below as Fig. 7.

One can see the wide distribution of error for the first data set; roughly 67% of the samples have magnitudes of errors less than 6mm. The next set of sample data (upper right of Fig. 15) quantifies the error distribution of samples that have been averaged over 10 points.

The standard deviation of this sample set was calculated to be 1.9mm. Likewise, the third and fourth sets of random data (lower left and lower right of Fig. 7, respectively) correspond to smooth surfaces that are defined by 100 and 1,000 LiDAR points, respectively. The corresponding standard deviations of these data sets were computed to be 0.6mm and 0.19mm, respectively. Resultantly, when scans are processed accordingly, the accuracy can actually be better than the error tolerances published for the scan hardware. Beam divergence also plays a role in the ability to acquire high detail at far distances and could be explored as an additional source of error; however, an analysis of beam divergence is beyond the scope of this paper.

Although it may not practical to scan a scene or surface 1,000 times, the principle holds true that the data converge to the true values as the number of samples increases. This principle can be utilized to recover information that exists below the noise level of the hardware to further optimize a data set.

## V. CONCLUSIONS

An error analysis of polygonal models created from actual and modeled terrestrial LIDAR data was performed. Ray traced points were compared to actual LIDAR scan data and modeled data using Fourier spectra, Taylor expansion and statistics. The standard deviation of error for the polygonal model of the anomalous groove was 0.474mm; the error for the bump was 0.599mm; and the error for the incline was 0.008mm.

This analysis showed that for objects with few high frequency features, scanning at a higher density actually increases measurement error. Polygonal models can act as filters to remove high frequency noise, thus lowering the errors inherent in LIDAR data utilization. Polygonal models created from LIDAR scan data can thereby be used to accurately represent real world objects. Furthermore, due to the error inherent in LIDAR scanning hardware, polygonal models can actually yield lower standard deviations of error than LIDAR scanning at a higher resolution.

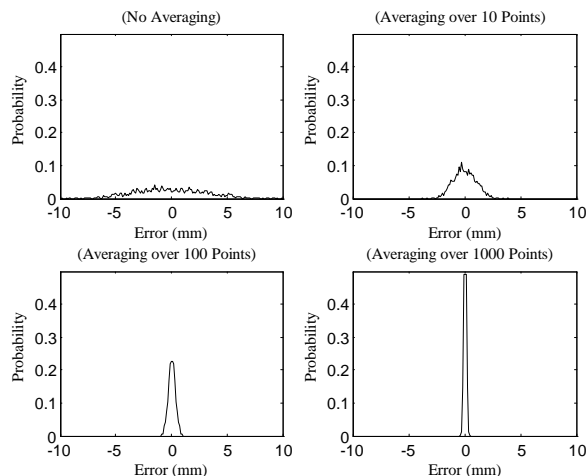


Figure 7. Probabilistic Curves Illustrating the Error of Scan Data

## REFERENCES

- [1] Smith, S.L., Holland, D.A., Longley, P.A., "The importance of understanding error in lidar digital elevation models", XXth ISPRS Congress, 2004.
- [2] Declercq, F.A.N (1996) Interpolation methods for scattered sample data: Accuracy, spatial patterns, processing time. *Cartography and Geographical Information Systems* 23(3), 128-144.
- [3] Desmet, P.J.J (1997) Effects of interpolation errors on the analysis of dems. *Earth Surface Processes and Landforms* 22, 128-144.
- [4] Maas, H-G and Vosselman, G (1999) Two Algorithms for Extracting Building Models from Raw Laser Altimetry Data. *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (1999) 153-163.
- [5] Morgan, M and Habib, A (2002) Interpolation of LiDAR Data and Automotive Building Extraction. *ACSM-ASPRS 2002 Annual Conference Proceedings*.
- [6] Rees, W.G. (2000) The accuracy of Digital Elevation Models interpolated to higher resolutions. *International Journal of Remote Sensing*, vol. 21, no 1, 7-20.
- [7] Sandwell, David T., "Biharmonic Spline Interpolation of GEOS-3 and SEASAT Altimeter Data", *Geophysical Research Letters*, 2, 139-142, 1987.
- [8] Smith, S.L, Holland, D.A, and Longley, P.A "Interpreting Interpolation: The Pattern of Interpolation Errors in Digital Surface Models Derived from Laser Scanning Data. *Proceedings of GISRUK 2003*.
- [9] Hyun S. Lee, Nicolas H. Younan, "DTM Extraction of Lidar Returns Via Adaptive Processing", *IEEE TGARS*, Vol 41, No 9, Sept 2003.
- [10] Wenge Ni-Meister, David L. B. Jupp, and Ralph Dubayah, "Modeling Lidar Waveforms in Heterogeneous and Discrete Canopies", *IEEE Transactions on Geoscience and Remote Sensing*, vol 39, No. 9, Sept 2001.
- [11] Wong, H., and Andreas Antoniou, "One-Dimensional Signal Processing Techniques for Airborne Laser Bathymetry", *IEEE TGARS*, Vol 32, No. 1, Jan 1994.

**Brandon J. Baker** is a student member of IEEE. He received the BSEE and MSEE degrees from the University of Utah in 1998 and 2003, respectively. He is currently pursuing the Ph.D. degree at the University of Utah. Brandon is President and CEO of Pinpoint 3D Inc., a LiDAR data processing and visualization company in Salt Lake City, Utah. His current research interests include applied mathematics, geospatial modeling, Light Detection and Ranging (LiDAR) data analysis, and numerical inversion theory. He has held positions as faculty in the Department of Mathematics at the University of Utah and Brigham Young University, as a research scientist at LiveWire Test Labs, as an ASIC design engineer at Evans and Sutherland, and a project engineer at InteliSum. He is the inventor or co-inventor of 16 pending patents.