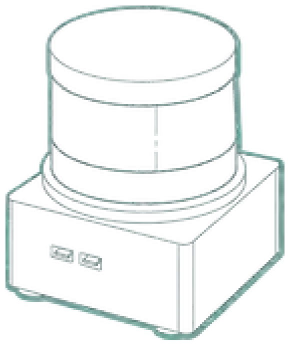


Measurement accuracy of Lidar-based SLAM systems

Kaarta Report 001

Abstract

We determine the accuracy of a lidar-based mapping system by determining differences in point cloud measurements. We use two methods to determine accuracy. The first method involves computing derived points and the second method compares point clouds by first aligning the point clouds. The results are that the lidar-based SLAM system does very well in comparison to stationary measurements with a high-end lidar.

September, 2016

Kaarta
5001 Baum Blvd, Suite 430, Pittsburgh, PA 15213

Table Of Contents

Overview

What is the accuracy of a 3D mapping system? Manufacturers often state the accuracy of the system is the same as the sensor used to collect the data, but this is only part of the answer since subsequent calculations affect the answer. Even the geometry and surfaces of the local environment can affect the measurements.

Also, *accuracy* is often confused with *precision*. They aren't the same. In this paper, we examine the problem, define accuracy, and provide some answers.

Background

Real Earth has developed a mobile 3D mapping system that creates a map of its environment even while moving continuously. It creates the map using only the range measurements it acquires along the way and a low cost MEMS IMU. No GPS is required and there is no prior map or other environmental infrastructure required. The accuracy of our point cloud-based 3D mapping system is determined by both the accuracy and precision of the sensing device and by the software used to create the map. Since a global 3D map is computed by incorporating a sequence of local maps, any noise or inaccuracy in the data propagates into downstream calculations.

Figure 1: The Stencil product from Real Earth is on the left and the FARO Focus 3D X330 on the right. Relative size is not to scale.

The robustness of the process for merging partially overlapping point clouds depends on the presence of strong features such as edges in the environment to establish correspondences between the point clouds. If the sensor is moving while acquiring point clouds, the algorithms use sensor measurements to map the environment and locate the sensor within that map. A benefit of determining the sensor position is determining the outward direction of the surface normal. For the Real Earth products, every point measurement in the cloud is made with the sensor in a different position and orientation. This process for analyzing the range data to build a map and determine localization is known as simultaneous localization and mapping or SLAM.

Various calculations and assumptions about measurement error are included when creating the resultant map. Conditions affecting data from real world measurements include lighting, atmospheric effects, environment, surface materials and textures, and, in the case of lidar, laser beam divergence. Divergence forms a larger spot at greater distances. On one hand the resolution and precision of the instrument may be very good; on the other hand there is a vast amount of data, hundreds of thousands of points, which are low-pass filtered to mitigate the noise. The accuracy specification of the instrument is only a starting point in understanding the accuracy of the measurement process.

Accuracy Specifications and Factory Calibration

A lidar typically produces data in the format <angle><angle><range>. There may also be reflectance data and, in some cases, a choice of signal or first return. These are usually different. Factory calibration defines the location and orientation of an instrument coordinate system inside of the lidar sensor head. The direction and range to a target in that coordinate system is determined from encoder positions that define the laser beam direction. If the calibration is accurate, scanned flat surfaces will be measured to be flat, and we will see accurate measurements of distances between derived points such as corners, radii of curvature, distances between parallel surfaces, and the angles between flat surfaces, etc. However, range noise can propagate into fitted parameters such as surface orientation or surface curvature. This, in turn, leads to uncertainty in the result. The resolution needed to detect a feature is a separate topic.

Factory calibration is static, though the sensor may be moved between measurements. Since the Real Earth product processes a stream of range measurements while moving, the accuracy of the resultant map needs to be compared with ground truth in a meaningful way.

To measure the accuracy, we compared outputs of two lidar units. One unit is the Real Earth Stencil product and the other is a FARO unit. We used a FARO Focus3D X330 lidar to define ground truth. To assess accuracy, we compared derived-point to derived-point measurements, and then compared the point clouds generated by the two systems. Clearly the FARO, while an excellent instrument, is not true ground truth – it too has a measureable accuracy that the manufacturer provides.

Assessing Accuracy

The accuracy of any sensor is determined in comparison to ground truth. Importantly, with a point cloud system, specific feature points on a surface are not measured, but derived. In other words, the laser beam from a laser rangefinder or a ray from a camera pixel does not intersect the surface at a specifically targeted location. Thus, derived points are computed, not directly measured. Examples of derived points include the center of a sphere which is computed from data points collected from the surface of the sphere, or of the location of a corner which is computed from the intersection of three planes. One way to determine accuracy is to compare derived-point to derived-point measurements.

Another way to assess accuracy is to measure the shape of a surface with known geometry, and then compare the results against a model. Examples include the radius of a sphere or cylinder, or the flatness of a plane, or the relative orientations of two planes, or a CAD model. However, creating large high accuracy shapes of known geometry is very costly.

Point clouds can be compared with carefully surveyed points acquired using expensive and sophisticated equipment. Often the output from these devices is accepted as a reference standard, but these measurements also have errors and noise associated with them. However, these measurements can be used as ground truth if their accuracy is at least an order of magnitude better than the system under test.

Accuracy is often confused with precision, repeatability and other terms. It is important to define the terms when having a conversation. Measurement accuracy, by internationally accepted definition, is the *closeness of agreement between a measured quantity value and a true quantity value of a measurand.*[1] The tricky part here is the term 'true quantity' because it assumes ground truth. Precision is *the repeatability or reproducibility of the measurement*.

In the following discussion, ground truth is defined by a FARO Focus 3D X330, shown in Figure 1, which is also a point cloud system. This lidar system is positioned at fixed locations to scan its environment. As a result, it operates in its calibrated mode at each position it is placed. The high-density point cloud it acquires at one position is then registered and merged with the point cloud acquired at a neighboring position. Registration of the point clouds can specifically take advantage of hand-placed fiducial features in the environment to ensure high accuracy.

A Comparison of Two Lidar-based Systems

To assess accuracy, we measured a natural environment that included man-made structures. We made measurements using two different Lidar-based systems and compared the results. The Real Earth Stencil product incorporates a Velodyne VLP-16 Lidar system that simultaneously scans 16 rotating lines oriented 1.875^{o} apart while the user manually carries the unit while moving to acquire dense data in 3D. The Stencil is a mobile mapping system that produces 3D maps in real-time without post-processing. The Real Earth system needs no GPS, no infrastructure, or any prior maps to build a 3D model.

The second system, a FARO Focus3D X330, was used as the reference to define ground truth. It acquires data while scanning its environment from a static position using a rotating mechanism. This system is phase-based and, along with similar products from Reigl and Leica, produces a dense point cloud whose accuracy is considered to be of 'survey quality', which is generally defined to be 5mm or less.

Range noise of the FARO instrument using a low reflectivity target is 2.2mm at 25 meters. Measurement accuracy of the scanner used in the Stencil product is 3cm up to 100 meters. This is an order of magnitude less than the FARO scanner whose measurements are therefore considered the reference.

Acquiring the Point Clouds

Point cloud data was captured using the FARO device and Real Earth's Stencil system. The FARO system was tripod-mounted at five fixed positions to obtain five independent point clouds. The scans were

[1] JCGM 200:2012. International vocabulary of metrology: Basic and general concepts and associated terms (VIM). France: BIPM; 2012.

made from rotating 360° around a vertical axis. The point clouds were registered to each other in a common coordinate system using FARO's proprietary software with spherical markers placed in the scene for further analysis.

When using the hand-held Stencil system, the operator simply walked around the natural scene while collecting the point cloud. A sweep consists of a set of range data in sixteen planes passing through the origin of the instrument's coordinate system. The complication is that the user is continuously "painting" the scene both by moving horizontally and tilting the scanner vertically, to acquire a full scene of 3D data. The algorithms used for this take the motion into account through rapid feature tracking and are assisted by an integrated IMU in this matter. The IMU data is used to unwarp each lidar scan and to provide a notion of relative motion to the scan matching algorithms and the final accuracy is not based on the accuracy of the IMU itself.

The scanned area is of a cabin of approximately 60 square meters (10.5m X 5.5m), with a roof and walls. Figure 2 shows the scene generated from the point cloud data. The cabin was on relatively level ground surrounded by trees and brush. A spherical marker was located several meters away from each corner of the cabin. The appearance of the spherical markers provided a way to compare derived-point to derived point distances. In addition, since the markers existed in both data sets, they provided a way to align and compare the Real Earth and FARO point clouds. Rapidly and accurately aligning two point clouds was useful for manually editing both point clouds at the same time to select features of interest.

In our evaluation, we compared the point clouds generated by the hand-held Stencil, which was continuously moving, and a tripod-mounted FARO system which needed to be relocated to several positions. The scan time to acquire the point clouds was about 10 minutes for the Stencil and 1 hour for the FARO device.

The FARO point cloud is used as a reference. To compare point clouds, we measured distances from the Stencil point cloud to the FARO point cloud. This is analogous to measuring the distance from a point to a surface. Small surface patches are computed for local groups of points within the FARO point cloud, and distances are computed to these patches from individual points in the Stencil point cloud. See Figure 5. Since the location of Stencil relative to the surface is found, the errors are "signed". Errors outside the surface of the object are positive, while errors inside are negative.

Derived-Point to Derived-Point Measurements

One way to assess accuracy is to compare the distances between *derived* points such as the centers of spheres or the intersections of three planes. We did this for the spheres present in the scene, and the corners of the cabin at a fixed height. Figure 2 shows the locations of the spheres in the scene.

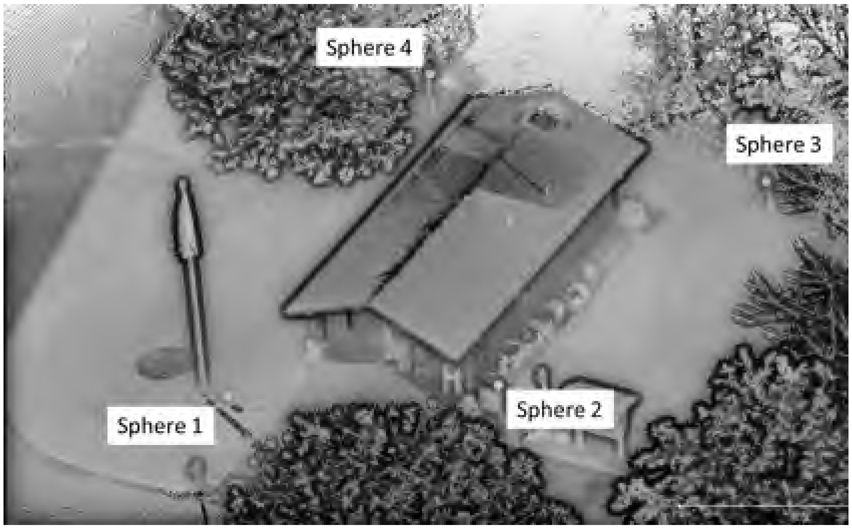


Figure 2: A view of scans from the Stencil and FARO systems.

We then aligned the point clouds to make the comparison easier. The point clouds from the Stencil and FARO instruments were roughly aligned in CloudCompare using an Iterative Closest Point (ICP) algorithm. ICP minimizes the mean square error. Then, man-made objects in the scene were kept while all of the foliage and the flag on the flagpole were cropped from the scene. We then aligned the remaining point cloud again. The alignment took place without rescaling since we assume the lidar sensors have been calibrated and range measurements are accurate. Our results and the alignment show this was a good assumption.

Figure 3 shows the remaining man-made objects in the scene. We kept several objects including the cabin, the flagpole, bulletin board, sign, flowerpots, spheres, and points from the side of a car in the parking lot. The trajectory of Stencil during scanning is shown as the red trace in the figure. Since only the measured surface coordinate data was provided, we used the closest Stencil data point (nearest neighbor) to each CloudCompare FARO based "core" point along with the Stencil position to determine the sign of the surface normal. The angle between the surface normal and vector from the surface to the sensor had to be less than 90 degrees; otherwise we flipped the direction of the surface normal.

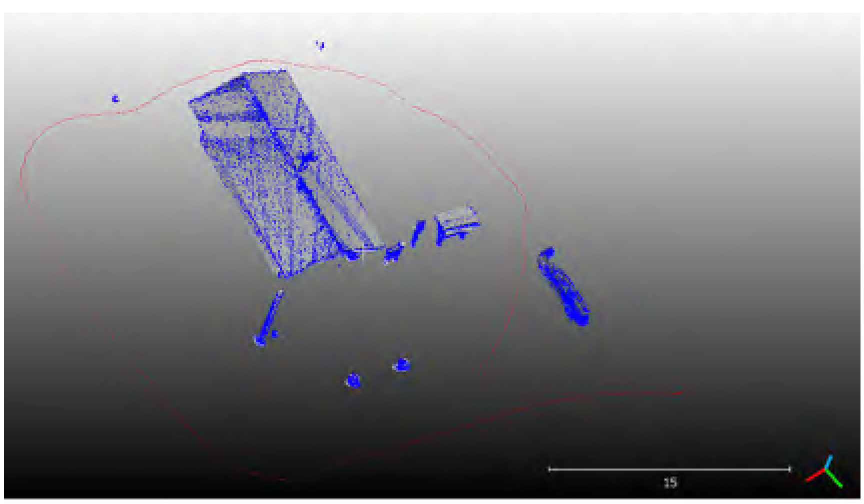
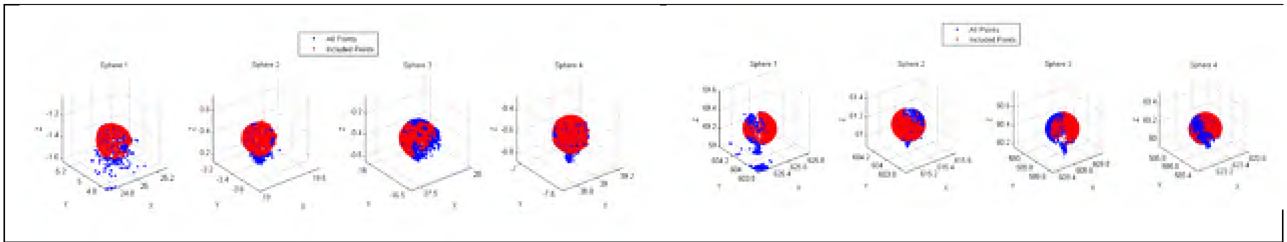


Figure 3: Man-made objects in the scene. The trajectory of the Stencil when measurements were taken is shown in red.

The direction of the outward surface normal in Figure 3 is in the general direction of the sensor. As a result, we know the sign of the error when comparing the point clouds. The merged point clouds for these objects were aligned and the resultant mean square distance between the clouds was determined to be less than 1 cm.

After aligning the point clouds, we identified the location of each spherical marker. We fit spheres to the local cloud of data to measure their center and radius. The spheres are lamp globes with a radius of 15.24mm (6 inches). They are translucent, so there is some uncertainty in their measured centroids and radii. Low pass filtering of the point cloud also affects these values. A measure of confidence in determining the locations of the sphere centers was the RMS (root mean square) error of the fit. Once the centers of the spheres were located we computed all of the mutual distances between the spheres.

Figure 4 shows the point cloud data in the vicinity of each sphere for both sets of data. Extraneous points from the ground or tripod along with instrument noise were suppressed by iteratively fitting a sphere to the data while removing outliers. That is, we fit a sphere to the data, removed outliers, and fit a sphere to the remaining points. The outliers are the blue points shown in the figures. The red points were included in the fit. A possible alternative is to use a random sample consensus (RANSAC) based algorithm. The result is the location of the center of each sphere, along with its radius. We computed the distances between the spheres using the computed sphere centers.



Stencil point cloud sphere data	FARO point cloud sphere data

Figure 4: Point cloud data from the spheres.

Table 1 shows the results. The differences between the Stencil and FARO measurements of the inter sphere distance is several millimeters.

	Stencil (meters)	FARO (meters)	Distance Error (millimeters)
Side 1-2	10.323	10.318	5
Side 2-3	15.349	15.351	-2
Side 3-4	14.291	14.277	14
Side 4-1	18.608	18.615	-7

Table 1: Distances between the spheres are shown in meters and the difference between the Stencil and FARO measurements are shown in millimeters.

The dimensions of the cabin were also derived from the point cloud data. The cabin is a classic structure with a gabled roof and four walls. The four corners of the cabin can be considered derived-points. The walls of the cabin are not perfectly flat and they include recessed features such as windows and doors. But these deviations are small relative to the size of the cabin and are measured by both systems. We fit a plane to each of the four walls of the cabin and then estimated the locations of the corners approximately one meter above the ground from both the Stencil and FARO point clouds. Figure 5 shows a rendering of the walls with the Stencil and FARO point clouds overlaid.

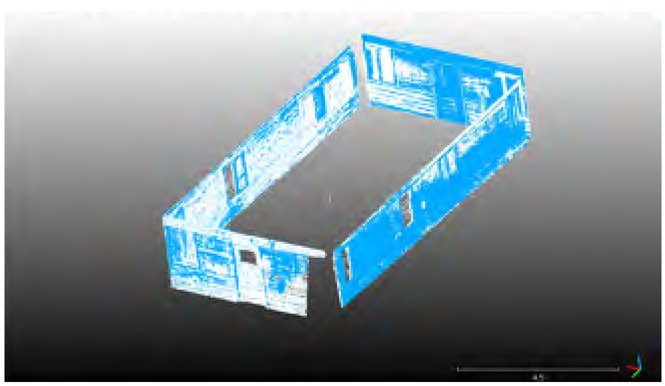


Figure 5: A rendering of just the walls of the cabin. We fit planes to the four walls and then computed a horizontal cross section to identify the four corner points of the cabin.

Figure 6 shows the coordinates of the four corners of the cabin along with the coordinates of the four spheres.

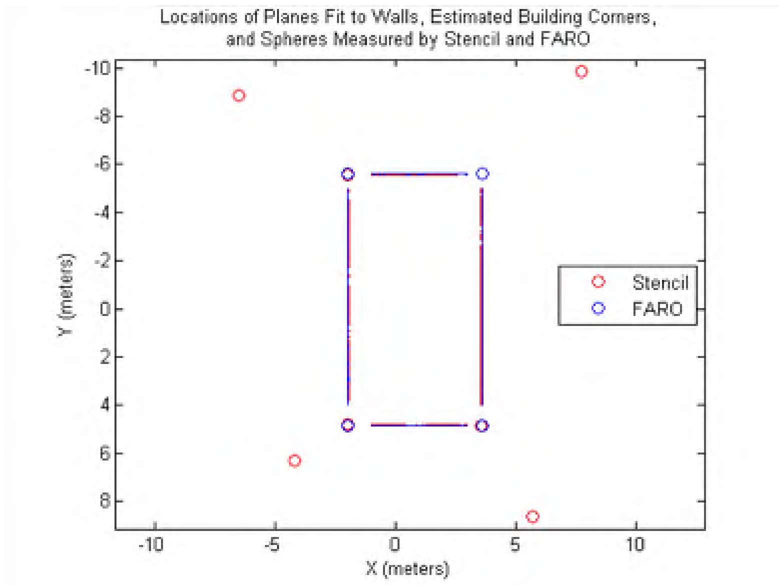


Figure 6: The locations of the four corners of the cabin are shown in the rendering. The wall lengths measured by the FARO are shown in Table 2.

Table 2 shows the estimated wall lengths measured one meter above the ground. The FARO estimates differ by 0.3% from the Stencil estimates.

	Stencil (meters)	FARO (meters)	Length Error (cm)
Wall 1	10.414	10.458	-4.4
Wall 2	5.538	5.554	-1.6
Wall 3	10.418	10.463	-4.5
Wall 4	5.553	5.563	-1.0

Table 2: Shown are the estimated differences in the wall lengths of the cabin based on the intersection of two adjacent walls. The third plane needed to define the four corner points is located approximately one meter above the ground.

Comparing Point Clouds

We then computed the average distance between point clouds using the FARO point cloud as the reference. To do this, a small surface patch is fit to local groups of points from the FARO point cloud. The distance from nearby points in the Stencil point cloud to each surface patch is then computed. The distances are signed since the surface normal points in the general direction of where the sensor was located when the data was acquired. See Figure 7.

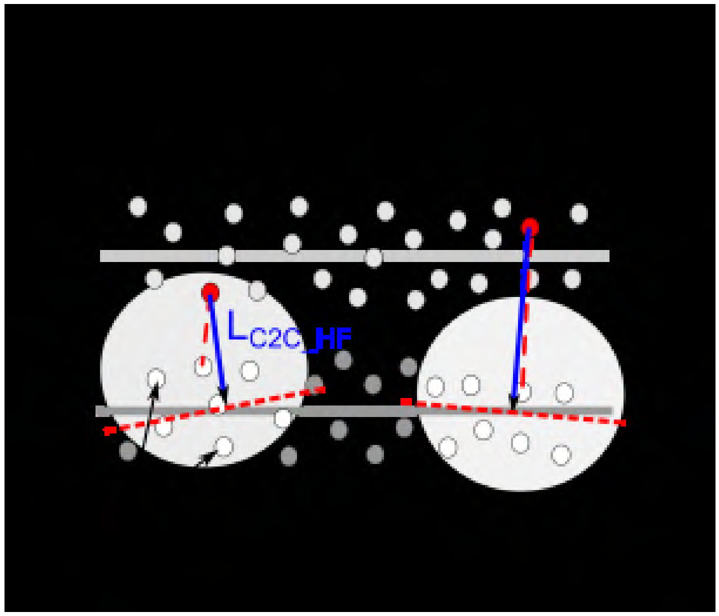


Figure 7: The graphic shows how distance is measured between the Stencil and FARO point clouds. Source: Cloud Compare M3C2 documentation.

A noise filter was applied to the point clouds from the scene with man-made structures and the point clouds were then aligned. The filter suppresses high frequencies associated with the sensor noise. It also rounds out sharp edges such as corners. Figure 8 shows a histogram of the errors. The shape of the histogram indicates there is a slight distortion of the Stencil point cloud relative to the FARO. The point clouds have been aligned to minimize the mean square error. Ideally, the errors seen in the histogram should be symmetric with an average of zero. When the point cloud shapes are in agreement the width of the histogram is proportional to the noise. The magnitude of the average error is approximately 7mm.

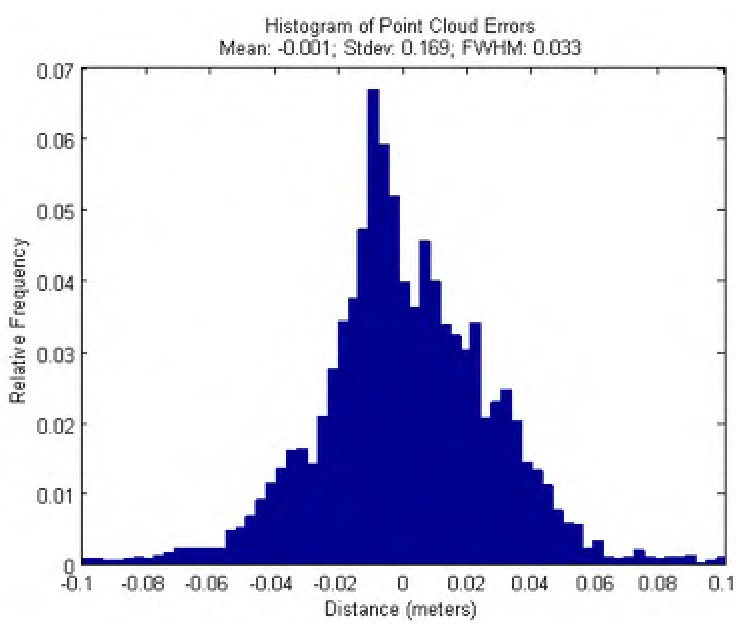


Figure 8: Histogram of point cloud errors for the man-made objects in the area. The caption shows the average error, the standard deviation, and the full width at half maximum of the histogram.

The point clouds for the cabin alone were then extracted from the data. The histogram of the errors was found to be bimodal as seen in Figure 9.

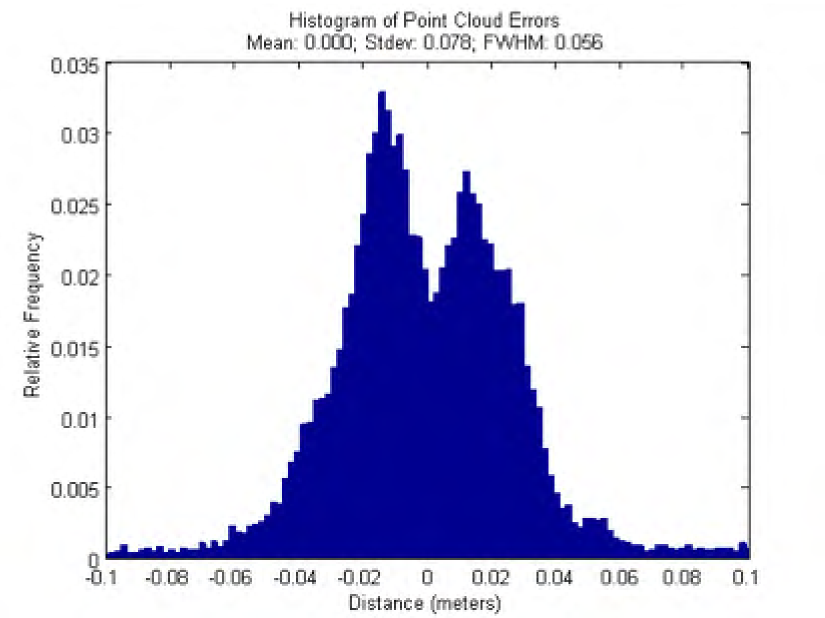


Figure 9: Histogram of point cloud errors in the data for the cabin alone. The caption shows the average error, the standard deviation, and the full width at half maximum of the histogram.

To understand the shape of the histogram we used a feature of CloudCompare to adjust the scale of data to improve the fit with the FARO point cloud. This scale change is a non-rigid transformation. Allowing the Stencil point cloud to be resized by 0.5% (a factor of 1.005) resulted in the uni-modal histogram shown in Figure 10. This scale change is equivalent to increasing every Stencil range measurement by 0.5%, suggesting a small error in calibration or point cloud alignment. The change of scale needed to improve the fit between the two point clouds is consistent with the differences in the lengths of the walls we computed using the derived corner points of the cabin. This is worth investigating further. Note that the scaling was used only for comparing the histograms. It was not used to improve the error differences between the FARO and Stencil cited earlier.

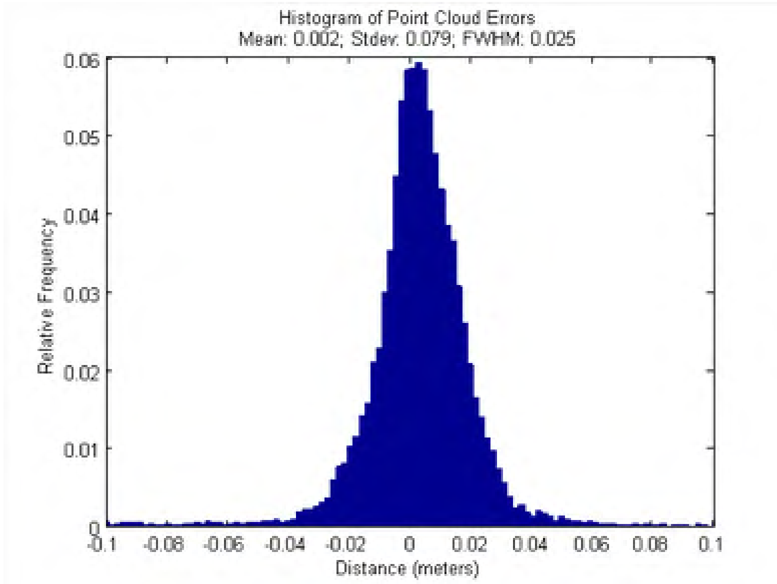


Figure 10: Histogram of point cloud errors in the cabin wall data using FARO as the reference after alignment of the cabin itself. The caption shows the average error, the standard deviation, and the full width at half maximum of the histogram.

The full width at half max (FWHM) of the histogram is 2.5 cm, or +/- 1.25cm. This is consistent with the noise specification of the laser scanner used in the Stencil after filtering. If there is an error in the measured shape of the cabin, not just scale, it should show up as a bimodal distribution in this histogram.

Summary

Real Earth's mapping system acquires and processes point cloud data in real time to assemble a point cloud map without the use of a GPS, maps, or infrastructure while it is in motion. The system was compared against a FARO Focus3D X330 system, which acquires point cloud data while stationary. The results show differences in the point clouds generated by the two systems on the order of 1 to 2 cm over an area of 20 square meters.

Differences in the point cloud measurements were determined in two ways. The first method involved computing derived points. The scene included spheres and a cabin with corners. Local point clouds associated with the spheres were extracted from the global data and fitting routines were used to estimate the centers of the spheres. The distances between the sphere centers as measured by the two devices were then compared. Corners of the cabin were determined from the intersection of planar surfaces computed from the cabin point clouds. Corner to corner distances were then compared.

The second method used to compare point clouds first aligns the point clouds. The point cloud acquired by the FARO instrument is considered the ground truth, since its *precision* is within one or two millimeters. We fit a surface patch to small groups of points acquired by the FARO and then compute distances to these patches from the point cloud. Ideally, the error histogram should be symmetric and the average distance close to zero. The width of the histogram, as measured by its full width at half maximum, is a measure of the random noise of the point cloud. Any break in the symmetry of the histogram indicates a problem merging of the point clouds that are used in creating the global point cloud.

In the example, point clouds were acquired in a natural environment with grass, trees and a flag moving in the wind, in addition to man-made objects. The results show that a mobile mapping solution can provide accurate information for architectural and engineering use. In summary:

- We collected point cloud data of a natural environment that included man-made structures. We used a FARO Focus3D X330 from fixed locations and Real Earth's Stencil system, which acquires point cloud data while moving.
- Point clouds acquired by the two systems were aligned using the open source program CloudCompare for comparison. Man-made objects including lamp globes, a cabin, and a vehicle were extracted from the scene to compare their point clouds.
- We processed the point clouds to compare distances between the spheres and the dimensions of the cabin.
- Analyses of point cloud errors indicate a small amount of distortion in the global point cloud. After local re-alignment error histograms are uni-modal and symmetric around zero error. Differences were within 0.2 - 1.4 cm in comparison with a stationary lidar for derived points at sphere centers. Differences were within 1-4.5cm for distances between derived points using wall intersections.

The overall accuracy of the resultant data was excellent and the mobile system can provide excellent information that can be used to model structures in CAD and provide direct measurements of buildings and structures.

References

[1] Velodyne VL0-16 Scanning Laser Range Finder Specification. www.velodynelidar.com.

[2] “Single-plane versus Three-plane Methods for Relative Range Error Evaluation of Medium-range 3D Imaging Systems”, David K. MacKinnon, National Research Council Canada, SPIE Proceedings, June 2015.

[3] “Towards the Development of a Documentary Standard for Derived-Point to Derived-Point Distance Performance Evaluation of Spherical Coordinate 3D Imaging Systems”, Bala Muralikrishnan et al., NIST. Procedia Manufacturing, V. XXX, 2015, pp1-12.

[4] “SLAM for Dummies”, Søren Riisgaard and Morten Rufus Blas. MIT ocw.mit.edu 16-412 online course.

[5] “LOAM: Lidar Odometry and Mapping in Real-time”, Zhang and Singh. Robotics: Science and Systems Conference, July 2014.

[6] CloudCompare v2.6.1 - User manual. http://www.danielgm.net/cc/documentation.html

[7] “Accurate 3D comparison of complex topography with terrestrial laser scanner: application to the Rangitikei canyon (N-Z)”, Dimitri Lague, Nicolas Brodu, Jérôme Leroux

[8] FARO Technology White Paper, Large-Volume 3D Laser Scanning Technology

[9] “Accuracy assessment of the FARO Focus3D and Leica HDS6100 panoramic type terrestrial laser scanners through point-based and plane-based user self-calibration”, Chow, Lichti, & Teskey, Canada. Laser Scanners II, 5865. Rome May 6-10, 2012.

[10] “Assessing the performance of 3D-imaging systems”, David K. MacKinnon. SPIE Newsroom. DOI: 10.1117/2.1201102.003393