Comparative Evaluation of Range Sensing Technologies for Underground Void Modeling

Uland Wong, Aaron Morris, Colin Lea, James Lee, Chuck Whittaker, Ben Garney and Red Whittaker

Abstract—**This paper compares a broad cross-section of range sensing technologies for underground void modeling. In this family of applications, a tunnel environment is incrementally mapped with range sensors from a mobile robot to recover scene geometry. Distinguishing contributions of this work include an unprecedented number of configurations evaluated utilizing common methodology and metrics as well as a significant in situ environmental component lacking in prior characterization work. Sensors are experimentally compared against both an ideal geometric target and in example void environments such as a mine and underground tunnel. Three natural groupings of sensors were identified from these results and performances were found to be strongly cost-correlated. While the results presented are specific to the experimental configurations tested, the generality of tunnel environments and the metrics of reconstruction are extensible to a spectrum of outdoor and surface applications.**

I. Introduction

RANGE sensing applications such as mapping, safeguarding and measurement are ubiquitous in robotics. Selection of sensors for these tasks is often made with complex considerations. However, comparative data between sensor models of the same class - and between differing classes - does not exist in any comprehensive form. Factory-listed accuracy values are established under ideal settings with ideal targets and methodology that are incompatible with other sensor calibrations. Moreover, the specifics of configuration and actuation also strongly affect the density and quality of data [1], [2]; configurations most useful in field robotics are often neglected in factory testing.

Comparative knowledge of sensor quality is particularly important when range sensors are used to create models for survey, digitization or surface analysis. Defining examples in this category include void mapping, mine-accident/event survey, and structural evaluation, where model accuracy directly governs issues of safety and culpability [3], [4]. The quality of reconstructed models is top priority even at the expense of system robustness or acquisition time. While mobile robots are used to generate void models, use of a stop-and-scan is acceptable. Features and statistics of the model - often volume and structural stability cues - are determined almost exclusively using post processing techniques.

The research described in this paper characterizes a broad spectrum of range sensors (Fig. 1) in the context of underground void modeling. Experimentation includes both a laboratory calibration utilizing ideal targets and a holistic comparison of models generated in representative underground environments. Both these comparisons utilize a common, repeatable methodology across all sensors and metrics of evaluation (sampling accuracy and density) that are inspired by the aforementioned real-world applications. It is noted that only the quality of geometric reconstruction from sensor data is evaluated. Cost, mass, energetics and field robustness are specifically not considered, as addressing the full gamut of issues would lead to intractability of experimentation. Moreover, factory values for these parameters are generally sufficient for decision making.

Foundations for this work were motivated by sensor selection in defense tunnel mapping application. While the primary results of this work are only strictly valid for the experimental configurations presented, the authors intend the work to contribute to shrewd sensor selection for a range of analogous environments and applications. The ideal component of the characterization is similar to much of the prior work in this area, with the exception that the current research addresses a much larger experimental set and is intended to be a broad basis for environment-neutral comparison. Generalizations such as intersections and corridors are extensible to most underground mines, caves and tunnels. Lastly, the emphasis on natural rocky materials is appropriate for many surface and planetary applications.

II. Prior Work

A rich body of work exists in range sensor characterization for the purpose of mobile robot mapping. The vast majority of prior work falls in the category of exhaustive single-sensor analysis, which this paper does not replicate in favor of a broader, holistic approach. The authors are not aware of any comparable research that has approached cross-sensor evaluation with the mix of methodology, scope and modeling relevance described in

Manuscript received March 27, 2011. This work was supported in part by the Robotics Technology Consortium, Award No. 69201002T04 and National Technology Transfer Center Award No. 0000007131.

Uland Wong, James Lee, Chuck Whittaker and Red Whittaker are with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.
{uyw, jslee1,warrenw, red}@andrew.cmu.edu

Aaron Morris is with AllPoint Systems, LLC.
aaron@allpointsystems.com

Colin Lea is with the mechanical engineering department at SUNY Buffalo, Buffalo, NY.
colinlea@buffalo.edu

Ben Garney is with Push Button Labs, LLC.
ben.garney@gmail.com

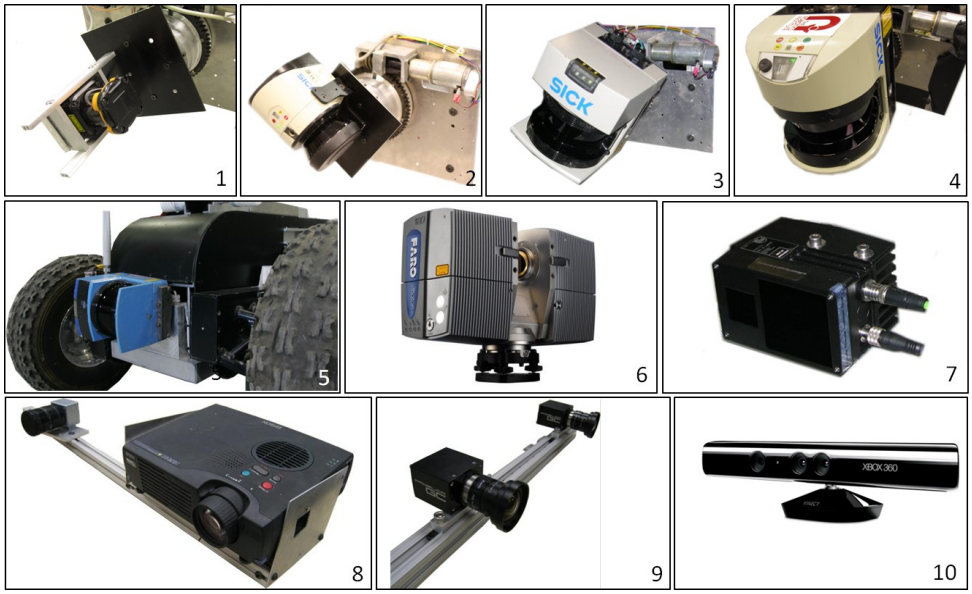

Fig. 1. Illustration of Sensors and Configurations Evaluated – (1) rotating Hokuyo UTM-30LX, (2) rotating SICK LMS111-10100, (3) rotating SICK LMS291-S14, (4) rotating SICK LMS511-10100, (5) rotating SICK LMS200-30106 affixed on a mobile robot, (6) Faro Photon80, (7) IFM O3D 201, (8) custom structured light sensor, (9) custom stereo vision sensor and (10) Microsoft Kinect. .

this paper.

Prior LIDAR characterization work is primarily either single-sensor physical analysis [5], [6], [7] or comparative evaluation between two sensors of the same class, one typically being a newer model [8], [9]. Rectangular targets of varying reflectance and orientation are commonly used for characterization. Though mixed-pixel analysis is performed in several of the studies, sensor capability in modeling rough unstructured surfaces or features smaller than the LIDAR spot size is not explicitly considered.

Conversely, prior work in evaluating vision-based sensors such as stereo and structured light have focused on the algorithmic aspects of enhancing image quality and feature matching for correspondence [10], [11]. This is perhaps due to the observation that measurement quality from triangulation sensors is most affected by the under-constrained nature of image formation, and not the physical implementation of the sensors.

Some prominent examples of related work comparing multiple range sensing modalities include RADAR and LIDAR for traffic negotiation [12]; photogrammetry and structured light [13]; LIDAR and stereo vision [14] and non-experimental comparison of LIDAR, triangulation, interferometry, and others [15]; and configurations for planar LIDAR actuation [2].

III. Characterized Sensors

Sensors were selected based on prevalence in robotics usage and availability to the authors (refer to Fig. 1 in Section 1 and Table 1); a total of 10 sensors were evaluated. Experimental configurations (i.e. actuation, physical parameters, and external illumination) were chosen to reflect optimality for void modeling at a critical sensing distance of 2-8 meters. This study is not intended to be a comprehensive sampling of sensor configuration parameters, but rather a broad sampling of sensor types specific to the application. For example, a baseline of 250mm and infinite focal distances were utilized for stereo vision; a less common configuration found in indoor robotics. There is no claim that results generated herein are strictly valid for any sensors or configurations other than those evaluated. A brief discussion of technologies sampled occurs in the following section.

Time of Flight - Time of Flight LIDAR sends a laser pulse in a narrow beam and measures the time taken for the pulse to return after reflection. Distance is calculated from the delay and knowledge of the constant speed of light. Most ToF lidars used in robotics application today are planar scanners in which a single point laser is rastered horizontally using a spinning right angle mirror. By rotating or nodding the scanner, a volume of 3D data is obtained. Multiple planar ToF sensors from SICK AG (LMS111, LMS200, LMS291, LMS511) and Hokuyo Ltd. (UTM-30LX) were characterized in this study to reflect the prevalence of LIDAR scanners in robotics. Three-dimensional data was collected by mounting to a 360 degree rotary actuator. This configuration was described by Omohundro as being optimal for void modeling as compared to simpler nodding configurations [1]. The rotary mount was actuated at a slow 1deg/sec to enable selective angular downsampling in post processing. As optical offsets are likely in mechanical actuation, proprietary dewarping software was utilized to estimate offset parameters.

Triangulation - Triangulating range sensors use the principle of intersecting geometric rays to measure distance to scene points. These rays must originate from two points

with fixed and known translation. Two types of vision-based triangulating sensors were evaluated: stereo vision and structured light.

Stereo cameras measure the distance to objects by comparing two images of a scene from two cameras and recognizing common features across the images. Triangulation of these "correspondences" with a known baseline produces depth. For testing, two Prosilica GC1290c 1.2MP cameras with 5mm lenses were parallel-mounted with a horizontal baseline of 250mm and calibrated. Lighting was provided in close proximity to (but behind the field of view of) the camera system during the underground modeling. Feature detection and point cloud generation was accomplished using the LIBELAS software for dense stereo matching [16].

Structured light sensors solve the correspondence problem by utilizing a light source (the dual of a camera) to unambiguously "paint" the scene with known identifiers. A single, static camera observes the change in brightness of scene points over time, which can then be decoded to generate correspondences between the source and camera. Two structured light systems were evaluated for testing: a custom sensor constructed by integrating a visible-light LCD projector with a Point Grey camera and the infrared Kinect™ sensor from Microsoft. The in-house solution utilizes a 40 degree field of view for both the camera and projector, which are angled 15 degrees inward with a 250mm baseline for optimal coverage and accuracy at two meters from the sensor. The Kinect was used in an off-the-shelf configuration with libfreenect to provide a 1-frame depth estimate.

Phase Shift - Phase shift, also known as FMCW (Frequency Modulated Continuous Wave), sensors continuously scan the scene with a periodic wave of light, modulating the frequency in a known, repeatable manner. The detected change in phase and frequency of the returned signal give the distance to the scene. This method of measurement produces high sample density and accuracy as there is no need to wait for a return pulse between measurements as in time-of-flight technology. A Faro Photon80 (FMCW LIDAR) was characterized for this study.

Focal Plane Array (Flash) LIDAR - Flash LIDAR operates like a digital camera, gathering information in a 2D array of pixels through lensing. Each pixel simultaneously measures a distance and reflective intensity each time a pulsed laser in the unit flash-illuminates the scene. The resulting snapshot provides a 3D model with reflectivity data. The IFM O3D201 sensor is a 50x62 pixel, 24fps flash LIDAR that was used in testing.

IV. Methodology

Methodology for comparative analysis of range sensors comprises two parts: (1) characterization in a lab utilizing an ideal, artificial target and (2) in situ characterization in representative underground environments. The purpose of the lab calibration is to establish a reliable baseline of comparison and to estimate uncertainty bounds for the sensors. As the true geometry of unstructured void environments cannot be known to arbitrary precision, the other sensors must be compared against the best performing sensor determined in lab trials, with uncertainty values informing a lower bound on error. Environmental characterization effectively stresses sensors with the gamut of materials, features and scenes likely to be found in the underground spaces of interest. Thus, in situ results enable a holistic view of performance that closely mirrors application intent. Both lab and in situ characterization are necessary to paint a complete picture of sensor quality.

TABLE I
EVALUATED SENSORS AND CONFIGURATIONS

Sensor Model	Technology	Evaluated Configuration
SICK LMS200-30106	Planar ToF LIDAR	0.5 x 180 degree rotating, 8m mode
SICK LMS291-S14	Planar ToF LIDAR	0.5 x 90 degree rotating, 8m mode
SICK LMS111-10100	Planar ToF LIDAR	0.25 x 270 degree rotating, 20m
SICK LMS511-10100	Planar ToF LIDAR	0.5 x 190 degree rotating, 24m clipped
Hokuyo UTM-30LX	Planar ToF LIDAR	0.25 x 270 degree rotating, 24m clipped
Structured Light*	Structured Light	PtGrey Scorpion w/ projector (1280x1024), 0.25m baseline
Microsoft Kinect	Structured Light	Off the shelf configuration, libfreenect, ~5m range
Stereo Vision*	Stereo Vision	2x Prosilica GC1290 (1290x960), ELAS, 0.25m baseline
IFM O3D 201	Flash LIDAR	Off the shelf, ~8m range
Faro Photon80	Phase-shift LIDAR	Off the shelf, 1/8 res, (0.072 x 360 x 320) degree, 24m clipped

*denotes an in-house implementation

A. Ideal Target Characterization

Laboratory calibration involves scanning a 1.25m x 1.25m, tiled and colored 3D checkerboard. While such "ideal" targets do not exist in field application, their artificial nature enables construction and knowledge of the true geometry to arbitrary tolerance. This information is useful in determining the true error of range sensors, which cannot be surmised in unstructured environments, as well as for testing the rare "edge cases" of sensor error. The checkerboard utilized is constructed to a tolerance of 1mm, beyond the expected single-shot accuracy of most contemporary range sensor technologies.

Features of the checkerboard are illustrated in Fig. 2. Two colors of semi-gloss dark tiles, raised 1.9cm (0.75"), are mounted on a neutral white diffuse backplane. Varying the surface reflectance of the tiles as well as the color enables characterization of sensor error as affected by target material. The pyramidal tiles located on the cardinal points of the target rise 3.8cm (1.5") from the backplane and are

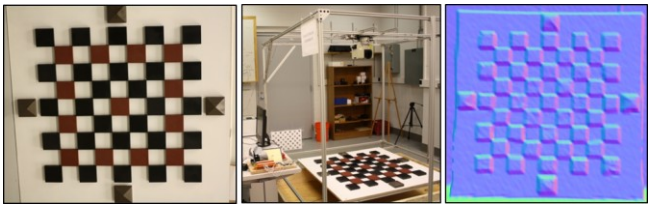

Fig. 2. A 3D checkerboard target used for ideal characterization (left), example experimental setup (LMS291 shown) for scanning the checkerboard (middle), and a mesh model of checkerboard generated using range data (right).

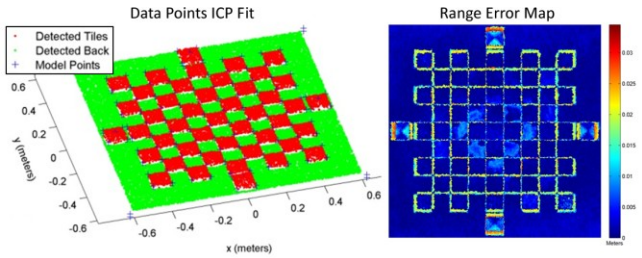

Fig. 3. ICP aligned range data from Photon80 with detected tiles and background (left) and Range error plot illustrating the "mixed pixel" effect near the edge of the tiles (right).

used to automate the process of aligning scans as well as to test pin-point sampling.

The target is centered such that the normal ray of the sensor passes through the middle tile. Scans are taken at a distance of 2.0m from the sensor origin and repeated for primary angles of 90 (normal), 67.5 and 45 degrees. Calibrated mount locations on a support frame provide ground truth for sensor positions from which the checkerboard is scanned (Fig. 2). Raw output from sensors is first transformed to point clouds with minimal filtering (no-return, max/min range). Point cloud data is then aligned with the ideal checkerboard model. While approximate sensor and target orientation are known, rotational ambiguities, inaccuracies in mounting and the intrinsic properties of the sensor result in error in raw data. Moreover, while the target may be oriented at a number of angles, the ideal model and error analysis assumes a fronto-parallelism. Utilizing initial estimates of sensor pose, the processing algorithm automatically detects the corner features of the checkerboard and finds a rigid transformation to the known model. A numerical optimization method, iterative closest point (ICP), is then used to fine-tune the alignment in the presence of non-rigid distortions and noise (Fig. 3). Points detected as the raised tiles are colored red, while points detected as part of the back plane are green. Corners of the files are marked with blue +'s. Quality metrics are computed after alignment.

Checkerboard data was collected for all 10 sensors. This data was evaluated using two metrics inspired by modeling applications, detailed below. Refer to Fig. 5 in the next section for illustration of these concepts on actual model data.

Range Error. The range error is the error between an observed data point and its known true location for a single measurement. The mean of the error distribution is a common measurement of accuracy. The range error used in this paper is calculated by aligning sensor data of the target to the ideal model using ICP and then raytracing the datapoints from the sensor origin. The L2-norm of the difference is the reported value. A large range error indicates an inaccurate or poorly calibrated sensor. The standard deviation of the range error is a measurement of precision.

Interpoint Distance. A frequent objective of 3D scanning is to create a mesh model or to infer surface geometry for object recognition. Both these applications require dense and regularly distributed surface samples. Interpoint statistics are generated by performing a 2D Delaunay triangulation on the surface points and measuring the distribution of resulting triangle side lengths. Large interpoint distances are indicative of "holes" in the model while a large variance in interpoint distances is indicative of badly shaped triangles. This statistic reflects the density of measurements on the target, which is an amalgam of angular density, sample rate, and field of view. Many actuated sensors which generate gratuitous readings but lack angular resolution in one or more axes exhibit inferior performance in resolving objects as compared to low-rate, fixed-resolution sensors.

B. Environmental Characterization

Three environments were selected for in situ mapping evaluation (Fig. 4): an unstructured corridor, an unstructured intersection and a closed structured tunnel. These selections were motivated by reason that generalizations of these environments represent the building blocks of most underground voids. For example, modern coal mines comprise rectangular grids of corridors and intersections while most caves, lava tubes and tunnels are simply lengthy corridors.

The "unstructured" environments mapped are located at the Bruceton Research Coal Mine in Pittsburgh, Pa. The corridor is a 60m tunnel with approximately 1.75 x 2.0m cross section, of which 20m was mapped with sensors. The intersection scene is a complex cross intersection of a wide corridor (3m) and a narrow corridor (1m). The primary construction in the mine environments is coal dust covered gunnite with a polyester mine curtain in the corridor and a wooden roof support in the intersection. The structured corridor is an indoor underground tunnel (formerly a mine tunnel), of 2.0 x 2.0m cross section at Carnegie Mellon University. It is made of a smooth concrete, painted white, and approximately 20m in length with a dead end. A series of heat and water pipes lines one of the walls.

The unstructured corridor was mapped by incrementing sensors longitudinally along a length of 20m. The number of scans (i.e. incremental distance) was calculated from the range and modality class of each sensor and ranged from 3, 7 or 20 scans. Blue fiducial cubes were placed in surveyed locations to enable fast stitching of incremental models, though the final alignments were tuned using iterative

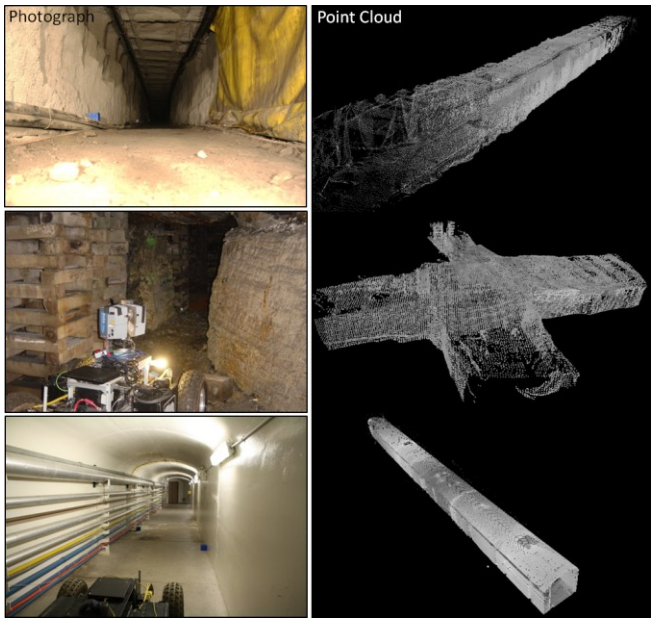

Fig. 4. Characterization environments from top to bottom: (1) unstructured corridor (2) unstructured intersection, and (3) structured corridor.

closest point (ICP). The intersection was mapped from a single static location, but sensors with narrow field of view were rotated to cover the scene. The structured corridor was mapped from a single location, aimed at the dead end. Dimensional and usage restrictions on some of the sensors prevented data collection from every sensor in all environments (ex. the LMS200 was affixed to a large mobile robot).

Metrics employed to assess range data collected from the sample environments are discussed below. These metrics are extensions of those used in ideal analysis with a few differences due to lack of known "ground truth" geometry of the environment. Chiefly, calculations are patch-based (quantized) to enable simple analysis on stitched models of multiple viewpoints without requiring 3D Delaunay and to establish approximation of true values from densely distributed local data. Where data density was insufficient for patch based analysis, less accurate aligned data from another sensor was used for comparison.

Range measurement sample density. Quality feature extraction from range data requires sufficient sample density. Sparse range data presents limited geospatial content and difficult determination of point associations. Moreover, robotic navigation processes such as obstacle detection, trajectory planning and other perception approaches rely on sufficient surface coverage to work properly. Surface coverage density is computed by counting the number of points that fall within two dimensional surface patches. The patches are approximately $100cm^2$ or 10cm x 10 cm in size (see Fig. 5). An upper threshold of 20 points (designated as "sufficient" coverage) was set, with the remaining cells containing 20 or less points. Thus, a curve of density vs. distance is generated. Range sensors typically display concentrated density in the near field,

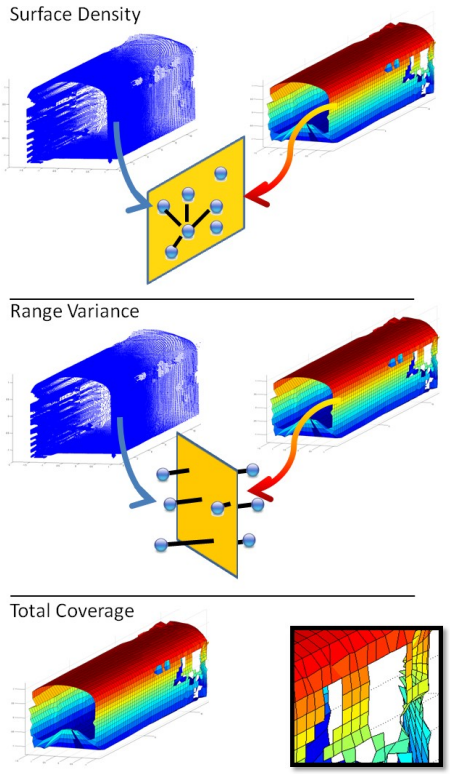

Fig. 5. Metrics of Comparison on Tunnel - (Top) Surface patch density is calculated from the number of points within proximity to a patch center. (Middle) Range accuracy is computed by comparison of data to a patch mean value. (Bottom) Holes in surface coverage are identified. Missing surfaces can be caused by occlusion, surface incidence angles or low reflectance.

which tapers off with increasing distance. Results summarized in Table II describe this metric in terms of Points-Per-Surface Patch (PPSP).

Range measurement accuracy over surface patches. Though dense, mean-centered samples may correctly approximate volumetric properties of the environment, high variance (excessive noise) results in "blurry" models, making minute features difficult to distinguish. Range variance is computed by applying a squared error statistic to the 10cm x 10cm surface patches. The error used is the normal distance from each measurement to the fitted plane for each patch. Results summarized in Table II describe this metric in terms of Variance-Per-Surface Patch (VPSP).

Amortized surface coverage. Long range feature identification can increase the accuracy and reliability of position estimation and fusion of multiple models. Short range or small field of view not only results in fewer distinguishing features for extraction, these features also have a smaller baseline, resulting in greater uncertainty. Moreover, short sensing horizons indirectly affect quality in automated model building by inducing erratic trajectory behavior. Lastly, complete surface coverage is required for accurate volumetric analysis of void models. Surface coverage is calculated by tallying patches that have been sampled. Patches associated with one or more points is labeled as "occupied." Patches with no points are labeled

"unoccupied." Surface coverage is therefore a ratio of occupied patches to a total patch count (determined manually from the size of the environment). This metric can be seen as a scalar summary of the density curve. Results summarized in Table II describe this metric in terms of Percent Surface Covered (PSC).

V. Results

Using the metrics described in the previous section, analysis was performed on sensor scans of the ideal target. The results are shown in Fig. 6. The x-axis (range error) is the empirical value of the accuracy and the y-axis (tri-neighbor interpoint distance) represents density. Sensors closer to the origin (zero) have better performance. The colored ellipses represent the uncertainty in the estimation of this value and are scaled by a factor of two for clarity. Experimental error, such deviations in mounting and data capture, as well as noise generated in the physical sensing process contribute to greater uncertainty.

The results show a natural grouping of the sensors into three performance classes. The Faro Photon80 was in a class of its own in regards to both metrics: a conclusion consistent with its pricepoint. As-built and survey LIDARs such as the Photon80 are designed to trade portability for maximal modeling performance.

All five planar time-of-flight sensors exhibited similar performance in a class below the Photon80, which is consistent with manufacturer specification and intended application (see Fig. 7). The LMS200, which has been a staple on underground modeling robots due to its lack of built-in filter, ties the LMS511 in accuracy and nominally wins out over the others. It should be noted that software issues prevented the LMS511 from operating at the highest angular resolution, though accuracy was unaffected. Had the sensor been capable of the factory maximum 0.125deg resolution, it likely would have been the best performing ToF LIDAR.

Inconsistent performers comprise the last class of range sensors. These sensors feature notable shortcomings in one or both of the metrics. Flash LIDAR is a nascent technology for outdoor sensing; the IFM O3D has comparable performance to the in-house designed stereo and structured light sensors, though it exhibits marginally better balanced performance and higher robustness. The structured light sensor has high range error arising from poor reflectivity that affects localization of light stripes at the highest scale, but the pattern and consistency of identified points is uniform and dense resulting in better target coverage. Stereo vision is strongly affected by the lack of texture and the repetitive tiling on the checkerboard. Depth estimation is generally accurate near the edges and corners of tiles and poor in the middle. The ELAS algorithm automatically rejects these ambiguous areas, leaving accurate points, but with large holes in between. Results from the Kinect sensor are intriguing. While the density score is skewed due to

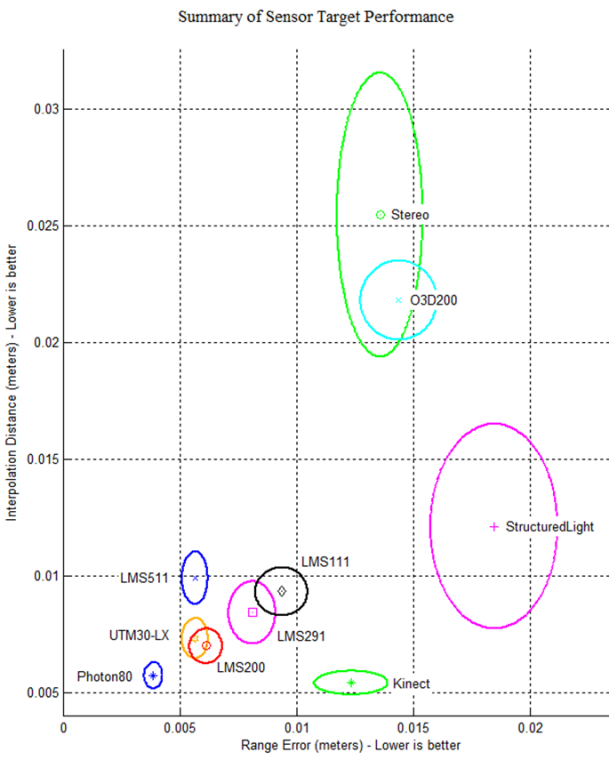

Fig. 6. Summary of Ideal Target Characterization

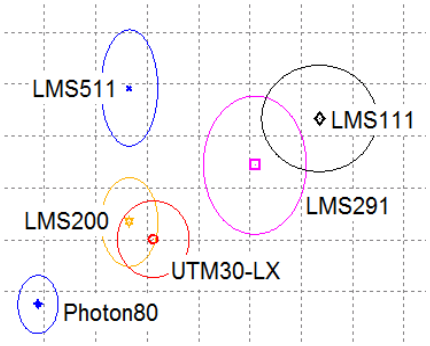

Fig. 7. Detail of Time-of-Flight Sensor Performance

fortuitous combination of narrow field of view and high density of the CCD, the pixel samples are not truly independent due to interpolation. However, the Kinect functions admirably as a low-cost volumetric mapper in this ideal case, greatly outperforming its pricepoint.

Analysis was also conducted in field environments to assess the impact of natural and artificial surfaces on each sensor. Each environment exhibits varying environmental materials, levels of occlusion and surface conditions that collectively affect performance of the range sensors. A montage of point clouds from several sensors in the structured corridor (Fig. 8) illustrates the wide disparity in scene appearance possible even from a single viewpoint.

Examined here (as described in Section 4.2) are the Points-Per-Surface Patch (PPSP), Variance-Per-Surface Patch (VPSP) and Percent Surface Covered (PSC). PPSP looks at the average number of points residing in a surface

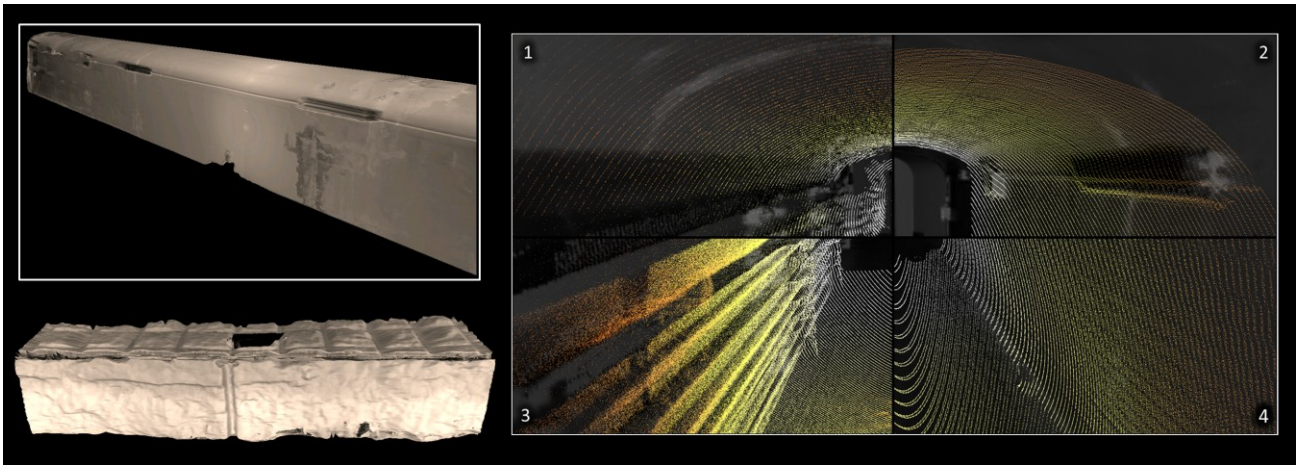

Fig. 8. Rendered Faro Photon80 mesh models from the structured (top) and unstructured corridor environments (bottom) are shown in the left image. A montage of point clouds (right) illustrates differences between the (1) Hokuyo UTM-30LX (2) Sick LMS291 (3) Sick LMS200 and (4) Sick LMS111 sensors. The points are mapped onto a grayscale albedo map of the structured corridor generated using the Photon80.

patch over all surface patches considered (spanning approximately 20m from the sensor). VPSP is the averaged variance of a surface patch over all surface patches considered. Lastly, PSC is the average Percent Surface Coverage, a ratio of the number of patches with data over the total number of patches that are averaged across all scans in a particular environment (see Table II).

The LMS111 came closest to the baseline sensor (Photon80) across all metrics. In fact, it exhibits a marginally lower measurement variance (VPSP) score than the Photon80 in several cases. Possible explanations for this anomaly include the presence of moving average filtering internal to the LMS111 (independent of the single-shot mode utilized) resulting in over-smoothed surfaces or a non-uniform angular bias of samples in each patch due to rotational actuation. The Photon VPSP should be taken as the best estimator to the true surface roughness value. The conclusion reached here is also contradictory to the ideal analysis, which nominally placed the LMS511 as the best performing time-of-flight sensor in accuracy. Generally speaking, the specific ordering of these sensors is mostly within the intra-class variance of ToF technology seen in this study. However, as density and accuracy estimation are not entirely independent, this effect could stem from lower numerical stability when fitting patch planes to the less angularly dense LMS511 data. The other time-of-flight sensors follow the LMS111 closely, with the particular LMS291-S14 unit generating anomalously high PPSP score due to a maximum range threshold of 8 meters. This threshold introduces a bias in regards to PPSP, which is a distribution-averaged value.

The PSC results arguably paint a more complete picture of density than PPSP. In this metric, the 8 meter range was a limitation for the LMS 291, while there was ample surface coverage by the LMS111 until approximately 20 meters. The Photon80 consistently captured the majority of the surfaces considered (with the maximum occurring in the unstructured corridor with approximately 52% of the

TABLE II
ENVIRONMENTAL SURFACE COMPARISON

	Unstructured Corridor			Unstructured Intersection			Structured Corridor		
	PPSP*	VPSP	PSC	PPSP	VPSP	PSC	PPSP	VPSP	PSC
Photon80	9.97	9.0e-4	52.08	8.34	0.09	30.05	9.16	3.1e-3	39.39
LMS111	18.58	2.0e-3	29.16	18.98	1.5e-2	23.71	18.49	1.0e-3	36.35
LMS200	-	-	-	16.52	0.12	24.75	14.18	9.0e-3	36.44
LMS291	16.98	5.2e-3	24.52	16.7	3.0e-3	18.83	17.27	4.9e-3	22.85
LMS511	19.10	6.2e-2	21.99	-	-	-	18.35	5.7e-2	42.30
UTM30	-	-	-	-	-	-	18.89	9.4e-3	33.96
O3D201	1.87	1.0e-3	3.14	2.04	3e-3	3.09	-	-	-
Stereo	19.58	7.31	0.45	18.00	1.7e-2	6.27	-	-	-
S. Light	-	-	-	-	-	-	12.84	4.9e-3	1.65
Kinect	13.01	8.7e-3	15.10	19.64	0.04	5.57	16.64	9.0e-3	11.44

*PPSP in units of points, VPSP in units of meters^2, PSC in units of percent.

available surface). Though the Photon80 is the baseline sensor, coverage values are not 100% due to holes and occlusions in the convex hull of measurement. In terms of relative surface coverage, the time-of-flight LIDARs achieved between 40-60% of the Photon80's coverage in the unstructured corridor, 60-75% coverage in the unstructured intersection and 75-100% in the structured corridor. These results correlate with the maximum possible sensing horizons (in decreasing order), where the photon performs particularly well in the lengthy unstructured corridor.

As a group, the inconsistent performers measured particularly poorly in density and coverage metrics in absolute terms, but did not underperform significantly in the accuracy (VPSP) metric. Once again the Kinect outperformed expectation with accuracy value about 50% less than the ToF LIDARs and coverage between 33-50%. It should be noted that while LIDAR sensors were thoroughly represented in this study, results for triangulation sensors could be considered data deficient. A multitude of possible configurations exist in camera and lens selection and matching algorithms. Intelligent enhancements to the evaluated setups could shift the results of triangulation sensors closer to LIDAR or decrease the uncertainty ellipses.

VI. Conclusion

This paper presents a cross-modal methodology and metrics for comparing range sensors in the context of underground void modeling. The modeling capability of 10 sensors - representing time-of-flight, triangulation, phase-shift and focal plane technologies - were evaluated on an idealized checkerboard target and in characteristic underground void environments. Three natural groupings of sensors were identified based on sample accuracy and density metrics. In general, sensors were found to be strongly cost-correlated, with a notable exception being the Microsoft Kinect which was found to greatly outperform cost. The authors expressly do not advocate for a particular "best" sensor, as this study considers only a sampling of possible parameters in the sensor selection function. The methodology and results presented are intended to supplement decisions for sensor selection and enable more accurate estimation of error bounds for similar applications. However, in the authors' motivating tunnel application, the Hokuyo UTM30-LX was found to have the right balance of performance, mass, features and cost.

Immediate roadmap for this work comprises the evaluation of additional sensors of interest. This includes greater sampling of stereo and structured light configurations and inclusion of omitted technologies such as multi-beam time-of-flight sensors (i.e. Velodyne HDL-64E) and commercial line scanners (i.e. Minolta 910). A database from this work is planned for public release to facilitate open involvement in the evaluation of new sensors. Possible future extensions include a more thorough sampling of experimental environments. Though this study has evaluated macro-scale geometry which represents the majority of underground environments, a variety of materials and surface geometry exists across tunnels, caves, mines and planetary craters. It is prudent to observe if or how performance varies with environment type and how closely the results presented in this paper hold for related environments.

Acknowledgment

The authors thank Dominic Jonak, David Wettergreen and Caterpillar Inc. for generosity of time and resources.

References

[1] Omohundro, Z.. Robot Configuration for Subterranean Modeling. PhD Dissertation. Robotics Institute, Carnegie Mellon University, 2007.

[2] Desai, A., Huber, D. Objective Evaluation of Scanning Ladar Configurations for Mobile Robots. In Proc. 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009.

[3] Morris, A., Ferguson, D., Omohundro, Z., et al. Recent Developments in Subterranean Robotics. Journal of Field Robotics, 2006.

[4] Wong, U., Garney, G., Whittaker, W., Whittaker, R. Camera and LIDAR Fusion for Mapping of Actively Illuminated Subterranean Voids. In Proc. of International Confe-rence on Field and Service Robotics, 2009.

[5] Okubo, Y., Ye, C., Borenstein, J. Characterization of the Hokuro URG-04LX Laser Rangefinder for Mobile Robot Obstacle Negotiation. SPIE Defense, Security + Sensing, Unmanned Systems Technology Conference, 2009.

[6] Ye, C., Borenstein, J. Characterization of a 2-D Laser Scanner for Mobile Robot Ob-stacle Negotiation. IEEE International Conference on Robotics and Automation, 2002.

[7] Kneip, L., Tache, F., Caprari, G., Siegwart, R. Characterization of the compact Hokuro URG-04LX 2d laser range scanner. IEEE International Conference on Robotics and Automation, 2009.

[8] Anderson, D., Herman, H., Kelly, A. Experimental Characterization of Commercial Flash Ladar Devices. International Conference of Sensing and Technology, 2005.

[9] Chiabrando, F., Chiabrando, R., Piatti, D., Rinaudo, R. Sensors for 3D Imaging: Metric Evaluation and Calibration of a CCD/CMOS Time-of-Flight Camera. Sensors, 9(12):10080-10096, 2009.

[10] Anchini, R., Loguori, C., Paciello, V., Paolillo A. A Comparison Between Stereo-Vision Techniques for the Reconstruction of 3-D Coordinates of Objects. IEEE Trans-actions on Instrumentation and Measurement, Vol. 55, No. 5, 2006.

[11] Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R. A Comparison and Evalu-ation of Multi-View Stereo Reconstruction Algorithms. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2006.

[12] Fischer, J., Menon, A., Gorjestani, A., Shankwitz, C., Donath, M. Range Sensor Eval-uation for Use in Cooperate Intersection Collision Avoidance Systems. In IEEE Proc. Vehicular Networking Conference, 2009.

[13] Radosevic, G. Laser Scanning Versus Photogrammetry Combined with Manual Post-modeling in Stecak Digitization. In Proc. 14th Central European Seminar on Computer Graphics, 2010.

[14] Matthies, L., Grandjean, P. Stochastic Performance Modeling and Evaluation of Ob-stacle Detectability with Imaging Range Sensors. IEEE Transactions on Robotics and Automation, Vol. 10, 1994.

[15] Blais, F. Review of 20 Years of Range Sensor Development. Journal of Electronic Im-aging, 13(1): 231-240, January 2004.

[16] Geiger, A., Roser, M., Urtasun, R. Efficient Large-Scale Stereo Matching. In Proc. Asian Conference on Computer Vision, 2010.