

Lumenhancement: Exploiting Appearance for Planetary Modeling

Uland Y. Wong

CMU-RI-TR-12-12

*Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Robotics*

The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

April 2012

Thesis Committee:
William “Red” Whittaker (Chair)
Srinivasa Narasimhan
David Wettergreen
Larry Matthies, *Jet Propulsion Laboratory*

Abstract

Planetary environments are among the most hazardous, remote and uncharted in the solar system. They are also critical to the search for life, human exploration, resource extraction, infrastructure and science. These applications represent the prime unexploited opportunity for automated modeling, but robots are under-utilized for this purpose. There is urgent need to explore, document, and evaluate these spaces with robots and to do so in a superior and efficient manner beyond the state-of-the-art.

This thesis introduces Lumenenhancement: the use of active illumination and intensity imaging with optical domain knowledge to enhance geometric modeling. While planetary environments are among the most challenging for robots, they share unique appearance constraints that can be exploited for sensing. Their dark, uniform, rocky and physically constrained nature enables a variety of physics-based vision techniques which are not pertinent in other field environments. Synergistic integration of calibrated imagery with traditional range sensing results in models with increased accuracy, sample density and readability. By leveraging the prevalence of existing illumination – such as sunlight – and common imaging sensors along with post-processing capability, this work promises broad significance.

Contributions from this thesis extend the state-of-the-art in several ways. Future discussion is anchored by experimental characterization of the planetary domain for the material and geometric properties of appearance. Material reflectance characterization using gonireflectometry has created the first empirical BRDF database of planetary materials. Studies of surface geometry have resulted in the first expansive database of range sensor comparative performance. The correctness of common vision assumptions in this domain, implications to intensity image techniques, and relevance to other domains are addressed. Novel methods for range and image fusion are devised to enhance and optimize aspects of model quality in the context of these principles, including geometric super-resolution, image-directed optimal sampling, and material classification. New possibilities for visualizing lumenenhanced models are also presented. Finally, implementations on mobile mapping robots and field experimentation at a coal mine and moon-yard are documented.

Acknowledgements

I would like to recognize the committee: Red Whittaker, Dave Wettergreen, Srinivasa Narasimhan and Larry Matthies, for expertly guiding me through the thesis process. To my advisor Red, thank you for providing the freedom to pursue a broad - and at times complex - agenda and the motivation to finish it. You fostered and championed this work from the very beginning. Dave and Srinivas, my graduate experience has been uniquely blessed by your generosity of equipment, lab space, advice and inspiration. I am undeserving of your immense support. Larry, your own vision research and review of this work are deeply appreciated.

I am fortunate to have had the help of my colleagues in the Subterranean Robotics group, both old and new, including Chuck Whittaker, Aaron Morris, Joshua Anhalt, James Lee, George Kantor, and interns Colin Lea, Chris Clingerman and Aashish Sheshadri among others. In particular, I would like to thank Chuck for driving, lifting, cutting, drilling and all the other things that were beneath his station to make my research go as smoothly as possible over the course of five years; I owe him a profound debt of gratitude. Aaron, you have been both a great role model and unyielding collaborator during my time here.

I acknowledge collaborators Ben Garney, Heather Jones, and Kevin Peterson. Ben, thanks for keeping me honest and providing much needed assistance.

A shout out to my Robotics Institute friends: Pras, Sam K., the original Masters of Disaster (John and Andrew), Jeff(s), Mike F., Young-Woo and others for supplying endless entertainment and indulging my delusions of grandeur.

Lastly, thanks to my mom for support over the years.

Table of Contents

Chapter 1: Introduction to Lumenhancement	19
1.1 The Case for Planetary Environments	20
1.2 Shortfalls in Planetary Modeling.....	22
1.3 Document Overview.....	24
Chapter 2: Foundations of Appearance in Modeling	25
2.1 Optical Sensing in Modeling.....	25
2.2 Domains as Appearance Spaces.....	28
2.3 The Role of Domains.....	36
Chapter 3: Thesis Statement	41
3.1 A Model-Centric Approach.....	43
Chapter 4: Background for Topical Areas	45
4.1 Planetary Modeling.....	45
4.2 Range and Intensity Fusion	48
4.3 Active Illumination	53
4.4 Range Sensing Technologies.....	60
Chapter 5: Characterization of Planetary Appearance	65
5.1 Analysis of Material Properties.....	65
5.1.1 Introduction to BRDFs.....	67
5.1.2 Methodology for Goniorelectometry.....	69
5.1.3 Data-Fitting Analytic BRDFs.....	77
5.1.4 Experimental Results for Materials.....	90
5.1.5 Discussion.....	107
5.2 Experimental Environments and Distributions of Materials.....	111
5.2.1 Variegation.....	118
5.3 Geometric Characterization	122

5.3.1	Barrenness.....	122
5.3.2	Accuracy of Modeling.....	128
5.4	Discussion.....	135
Chapter 6: Planetary Case Studies for Lumenhancement.....		137
6.1	Camera and LIDAR Fusion for Super-Resolution Modeling.....	137
6.1.1	Foundational Work.....	139
6.1.2	Fusion in the Markov Random Field Framework.....	139
6.1.3	Experimental Results.....	145
6.1.4	Analysis	151
6.1.5	Discussion.....	153
6.2	Image-Directed Sampling for Geometric Modeling of Lunar Terrain	154
6.2.1	Foundational Work.....	156
6.2.2	Frequency-Based Modeling.....	157
6.2.3	Modeling for Planetary Robots.....	163
6.2.4	A Class of Image Directed Scanners	170
6.2.5	Discussion.....	175
6.3	Utilizing the Sun as a Motional Source	177
6.3.1	Super-Resolution Revisited	177
6.3.2	Material Classification using Goniorelectometry In Situ	182
Chapter 7: Conclusion.....		189
7.1	Summary of Results.....	190
7.2	Contributions.....	191
7.3	Impact.....	192
7.4	Future Work.....	193
References.....		195
Appendix.....		203
7.4.1	BRDF Polar Plots.....	203

Table of Figures

Figure 1. Robotic Modeling in Planetary Environments. Examples include (1) automated lunar landing, (2) terrestrial underground inspection, (3) lunar resource extraction, (4) mapping of “skylights”, (5) exploration of abandoned mines. [Photo courtesy Red Whittaker, CMU for (1); David Wettergreen, CMU for (2); all others Uland Wong].	21
Figure 2. Canonical Optical Sensors. Cameras measure a composition of material, geometry, and illumination using the mechanism of two dimensional image projection (top). Only angle and intensity information for each light ray is preserved. Range sensors directly measure geometry to produce “clouds” of point samples in three dimensions (bottom). Data is from a sandstone cave.	27
Figure 3. Modeling Robots and Their Sensors. Some examples of field robots utilizing LIDAR ranging (yellow arrows) and cameras (green arrow) in varying configurations. (1) Cavecrawler, a mobile mine inspection robot; (2) Ferret, a snake-like borehole inspection robot; (3) Scarab, a lunar prospecting robot [photo courtesy David Wettergreen, CMU]; (4) an automated lunar lander concept; (5) Depth-X, an underwater explorer that utilizes sonar and a camera [photo courtesy David Wettergreen, CMU]; (6) H1ghlander, an off-road racing robot [photo courtesy Red Whittaker, CMU].	28
Figure 4. Surface Appearance Model. Incoming light rays, parameterized by polar angles, strike the object surface. These rays are decomposed and reflected in many directions determined by the material properties. Some of the outgoing rays will reach the sensor location indicated by the viewing direction.	29
Figure 5. Parameters of Image Formation can be divided into two classes: those intrinsic to the application environment (material properties, surface geometry and natural illumination) and those which are byproduct of sensor design.	31
Figure 6. Mondrian Environments describe the simplest conditions of image formation which are matte materials, ambient illumination and planar surfaces.	32
Figure 7. The Planetary Domain and Related Environments. The axes denote increasing quantity and complexity of materials; geometry, which includes macro-scale features and minute surface detail; and illumination – number and complexity of natural light sources. The origin represents a perfect Mondrian environment. Other environments shown include: (1) Lunar and terrestrial lava tubes, (2) underground mines, (3) Lunar craters, (4) asteroids, (5) the Martian surface, (6) the sea floor and underwater caves. Work presented in this thesis applies to (1-4 in red), while the subterranean domain of Omohundro and Morris is illustrated with the red oval.	33
Figure 8. Sample Artificial Light Source Configurations. A simulated rocky scene is illuminated by (1) a single near point source, (2) a directional linear source, (3) multiple near point sources, (4) multiple, spectrally-distinct sources.	55
Figure 9. Role of the BRDF in the Geometric Appearance Model. The BRDF is a material-specific mapping of reflected light based on the direction of incoming and outgoing rays.	66
Figure 10. 3-DOF Gonioreflectometer Configurations. A traditional actuated setup where a single light source can move about the hemisphere of ray directions, but the detector is constrained to an arc with static azimuth (left). A redundant static setup developed in this work which utilizes multiple sources and detectors in the hemisphere; however, the cameras are co-planar (right).	69
Figure 11. Photo of cube gonioreflectometer illustrating major components: fixed lights, cameras, test sample and LIDAR for self-calibration.	70

Figure 12. CAD model of cube frame with ideal light source positions (left). Plot of incident light rays from source positions distributed approximately uniform across the possible range of elevations (right).71

Figure 13. Detailed view of gonireflectometer components. (Left) digital relays toggle power to light sources, (middle) each illumination source consists of a tight cluster of 3 LEDs each with their own hemispherical diffusers and (right) SLR camera and LIDAR mount positions at the top of the frame.....72

Figure 14. Geometric calibration utilizing an “as-built” survey scanner (left) and model of cube showing planned (blue) and as-built (red) positions of the light sources (right).....74

Figure 15. Calibration of incident irradiance using an ideal Spectralon sample (left) and compensation for distance falloff from light source positions on cube frame (right). The sources are small enough that the incident light on the target does not deviate significantly from ideal.....76

Figure 16. Canonical illustrations of the five BRDFs analyzed: (left to right) Lambertian, Oren-Nayar, Phong, Cook-Torrance-Sparrow and Hapke. They are rendered on the Eros model with straight on (top row) and glancing (bottom row) illuminant directions.79

Figure 17. Microfacet Surface model utilized in Oren Nayar and Torrance BRDF models. Surfaces are randomly oriented collections of microscopic facets arranged in symmetric v-shaped cavities (left). Light reflecting from the source to the viewer is attenuated by the surface geometry by shadowing, where the source is blocked (middle) and masking, where the viewer is blocked (right).82

Figure 18. Macbeth Chart (left) used for color calibration and estimated whitepoint for each light source (right).90

Figure 19. (Left) Design of CMU-1 lunar regolith simulant. The mixture ratio of coal and limestone dust is found by color and albedo matching against a known target under ambient illumination. CMU-1 (smaller-right) is qualitatively very similar to the fine portion of JSC-1A (larger-right), a NASA developed simulant.93

Figure 20. CMU-1 BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.....93

Figure 21. Coal Dust BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.....95

Figure 22. Gunitite BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.....97

Figure 23. Granite BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.....98

Figure 24. Sandstone BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.99

Figure 25. JSC-1A BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue..... 101

Figure 26. Limestone BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue. 103

Figure 27. Ochre Paint BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.	104
Figure 28. Spectralon BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue. There is bias error for high-intensity measurements due to semi-saturation. This and other errors prompt regularization of calibration parameters, regardless of “known” sample values.....	105
Figure 29. Rendering of Fitted Planetary and Reference Material BRDFs. Color, relative albedo, and reflectance functions are derived from experimental data; brightness is enhanced for print and screen viewing.....	109
Figure 30. Typical Imagery from environments considered in characterization: (left) Bruce-ton Coal Mine, (center) macroscopic Lunar terrain and (right) Walker’s Mill sandstone cave.	111
Figure 31. Approximate Prevalence of Surface Materials in Experimental Environments from qualitative analysis. Notes: the other category of Bruce-ton mine denotes artificial ceiling materials and mining equipment. Gunite+coal denotes a discrete variegated surface of these two materials. The rock+regolith category under the moonyard column denotes a variegated surface of regolith and rock.....	111
Figure 32. Overhead Map of Bruce-ton Mine, the primary experimentation environment for mobile robots in this thesis. This mine consists of two appearance environments, (1) Corridors consisting of Gunnite covered walls and ceiling and dirt floors which constitute over 98% of the surface area of the mine and (2) open coal faces which are less than 1%. Internal objects, such as mining vehicles constitute the remainder of surface area. Blue areas represent the open void of the mine within the solid strata.	112
Figure 33. 3D mesh model of corridor in Bruce-ton Mine, generated by mobile robot.....	113
Figure 34. A simulated lunar rock distribution utilizing Surveyor data (top left) and realization in the construction of an artificial moon scene (top right). The bottom row illustrates a z-colored mesh model of the scene generated with LIDAR scanning (bottom left) and color-mapping of the terrain from DSLR images (bottom right).....	115
Figure 35. Distribution of Rock Sizes (left) at the Lunar Surveyor landing sites, from NASA Surveyor Project Final Report 1968. Creation of the moonyard utilized crushed and sieved limestone rocks of different diameter (right).	116
Figure 36. A partial LIDAR scan (left) and photo (right) of Walker’s Mill cave showing the irregular geometry of fractured sandstone.	117
Figure 37. Variegation of Materials in Environments. The walls of Bruce-ton Mine (top row) show variegation between gunite and coal, and occasional artificial objects. Lunar terrain (center) was only numerically analyzed for the moonyard, which is an artificial average case. Variegation is believed to be very low. Walker’s Mill Cave (bottom) shows anisotropic banding of the sandstone material, but variegation of material is very low. The right most image shows a single example of plant roots.....	118
Figure 38. Variegation Estimation from Images. A color image taken under near-ambient illumination (left) is clustered with the Mean-shift algorithm using high noise tolerance and the $L^*a^*b^*$. A few large clusters are labeled annotated (right). The number and size of the resulting clusters determines the variegation of the scene.	119
Figure 39. Variegation in Bruce-ton Mine Images. The uniformity score (blue) and the scaled, inverse number of clusters (red) are plotted on the same axes for comparison. Higher numbers represent more uniform scenes.....	120

Figure 40. Variegation of Indoor Scene. This garage scene required 367 clusters, and produced a uniformity score of 0.0096, less than half that of Bruceton mine.....	121
Figure 41. Correlation of Image Intensity to Surface Normals. An ultra-accurate geometric model (1) is used to produce surface normal estimates at each voxel location via differentiation (2). The slant angles calculated from these surface normals are correlated with pixel intensity from an ambient image (3) and a point-illuminated image (4) that satisfies the planetary assumptions.	122
Figure 42. Correlation of intensity values to surface gradients for the moonyard under two different lighting assumptions. This graph shows that simple point illumination promotes superior surface recovery. The optimal value occurs at n=32 for ambient and n=16 for simple lighting.	124
Figure 43. Correlation of intensity values to surface gradients in Bruceton Mine Imagery. A variegated scene with patches of gunitite and coal is compared against a uniform scene consisting only of gunitite. Uniform scenes improve the accuracy of geometric recovery with intensity images.	125
Figure 44. Archetypical Component Geometry in Coal Mines. Point clouds of corridors, 3-ways and 4-way intersections are utilized in testing how macroscopic geometry affects the smoothness constraint (left). The holes in the point cloud are from sensor self-occlusion, and do not effect computation. A barrel against the backdrop of a smooth wall creates an occlusion edge which strongly breaks the smoothness assumption (right).	126
Figure 45. Autocorrelation of Interpolated Normals from Mine geometry. Areal density at the maximum range of 7m is about 4 range samples per cm ² at native resolution (512x) down to about 0.18 samples per cm ² for 1x density. Data from three geometrically distinct scenes within the mine are shown: a straight corridor, a 3-way intersection and a 4-way intersection.....	127
Figure 46. Autocorrelation of Interpolated Normals from Moonyard geometry. Areal density ranges from 1mm grid size at 512x resolution to a 23mm grid size at 1x.	128
Figure 47. 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.	129
Figure 48. 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).	130
Figure 49. 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).	131
Figure 50. Summary of Ideal Target Characterization with detail of planar time-of-flight LIDAR performance (inset). All statistics are for single-shot measurement where applicable.....	133
Figure 51. Sparse LIDAR range readings(left) and high resolution intensity images (center) can be fused to create super-resolution models (right) with a Markov Random Field.	137
Figure 52. A mine inspection robot with active illumination, <i>CaveCrawler</i> , in process of corridor modeling (left) and raw fisheye photographic data (right). Note that the left image utilizes a different lighting configuration than presented here.	138
Figure 53. Markov Random Field Graphical Model. Green nodes (I) represent the image pixel data, brown nodes (x) represent the hidden true range value to be estimated, aqua nodes (R) represent the sparse range data and the blue node represents the interpolation uncertainty estimate. There is 1 pixel value for every hidden node (x), but there may be many nodes without a corresponding range value (R).	140
Figure 54. MRF Super Resolution Process. (1) Raw LIDAR point data is converted to a range image from the camera perspective. (2) Specularities are removed from the color HDR imagery to produce a	

diffuse image. (3) Surface normals are estimated from the diffuse image using shape from shading. (4) The surface normals and the range image are fused in the MRF framework.....	144
Figure 55. Intermediate Representations of Data for Fusion. (1) Raw fish-eye Image of a mine scene. (2) Range Image (depthmap) from raw LIDAR readings. Depthmap shown is warped to the space of the fish-eye image. (3) Ground truth depthmap. (4) Irradiance compensated color image, dipped to the boundaries of the LIDAR data. (5) Intepolation Uncertainty map. White values indicate scan points, while varying degrees of gray indicate increasing interpolation distance between scan points. (6) Specular-Invariant image after SUV transform. (7) Shading estimate from intensity image. (8) Surface Normal map from shading estimate utilized in MRF. RGB channels correspond to magnitude in XYZ Cartesian coordinates of unit normal vector. (9) Super-Resolution point cloud generated using MRF technique, showing detail of roof supports (right side of image).....	145
Figure 56. (Left) Experimental setup with 1. LIDAR scanner. 2. Fisheye Camera, 3. Light Source. (Center) Raw fisheye imagery. (Right) Ground truth range image.....	146
Figure 57. Reconstruction Improvement vs. Raw Interpolation.....	147
Figure 58. Point Cloud of Cribbing. Low resolution cloud (left) and high resolution reconstruction from algorithm (right) showing stacked timbers supporting the roof.....	148
Figure 59. Colorized 3D Reconstruction. Full scene (left) and mine curtain detail (inset and right).	148
Figure 60. Mine Corridor 3D model. (1) External view. (2) Internal view with rail tracks.	150
Figure 61. Detail of Hole Filling Process. Dense underlying geometric points are connected with their neighbors using multi-scale texture. Surface normal estimates from image data provide discontinuity checks and perception of shape.....	151
Figure 62. (Left to right) Roof supports covered in Shotcrete; Raw intensity to ground truth correlation; Shape-from-Shading estimates to ground truth correlation; and Reconstruction error reduction. Scale is brown to white over [-1, 1] for correlation and navy blue (0.05m error reduction) to red (0.05m error increase). Discontinuity edges present the great amount of error increase.....	152
Figure 63. Scenes with Poor Results. Irradiance compensation breaks down when the robot approaches too near a wall (left). A large open area is not sufficiently illuminated by the robot's light source (right).	153
Figure 64. Approximate Spatial Uniformity from Adaptive Sampling.	154
Figure 65. Automated Lunar landing is one possible application of image-directed sampling. A robotic lander (left) must model the near-field terrain to select the best landing site closest to a location of interest. Images generated during descent (right) can guide scanning with a laser altimeter to resolve regions of ambiguity. [Lander photo courtesy Red Whittaker, CMU; Apollo image AS15-M-0103 from NASA/JSC/ASU].....	155
Figure 66. Example Spatial Sampling Frequency Distributions based on application intent.	158
Figure 67. The frequency content of a region is proportional to local sampling density. This relationship is illustrated empirically by decomposition of a range image with wavelets (top), and reconstruction by downsampling a point cloud (bottom).	159
Figure 68. Image-Directed Sampling Process. (1) Intensity Image features generate a sampling probability map. This map approximates the true surface frequencies of the scene. (2) Optimal samples are selected using the probability map and acquired with a manually guided LIDAR scanner. (3) The process is iterated until stopping conditions are satisfied. (4) Triangulation of the raw LIDAR points produces a 3D reconstruction.....	162
Figure 69. Image features in diffuse environments, such as the moon, correlate strongly with geometric features. These three simulated lunar terrains are utilized in this work.....	163
Figure 70. Lunar Analog Terrain with simulated overhead sun light.....	167

Figure 71. Comparison of selected strategies over all subsample amounts and metrics on lunar terrain model. Visually, wavelets are the best performers (green).	168
Figure 72. Lunar terrain model generated with image-directed sampling. (1) An extremely sparse sampling rendered with roughness-modulated Oren-Nayar BRDF retains useful geometry and is visually appealing. (2) The frequency map of the terrain generated using wavelet decomposition. (3) Detail of roughness visualization demonstrates that a data deficient model can retain much of the optical appearance of the highest resolution model.....	169
Figure 73. A borehole-deployed underground inspection robot scans with a single beam LIDAR by actuating pan and tilt axes. Mechanical steering constrains the placement of the next sample based on the position of the prior sample and limits on rotational velocity. An independently control camera can image the scene with an LED light source.	170
Figure 74. Actuation Trajectories simulated on range images of a scene. (1) Linear scanning produces a coil of data parallel to the scene. (2) Raster scanning oscillates the sensor head on a single axis while maintaining motion on the orthogonal axis. (3) Adaptive spiral scanning oscillates both axes, varying the tightness of the inner spiral based on data collected along the outer spiral. (4) Independent sampling selects locations at random for acquisition. This may require erratic actuation or an independent sensing array.....	171
Figure 75. Iterative refinement can be utilized in motion-constrained scanning to enable variable density sampling in regions while respecting the grid nature of actuation (1). A mine scene is scanned with Image-Directed Sampling using the <i>Ferret</i> borehole robot (2). Uniform surface sampling of the mine scene (bottom) illustrates much less aliasing than traditional naïve scanning.....	172
Figure 76. An image-directed scanner consisting of a high resolution camera optically co-located with a DLP projector (left). Experimental setup utilized in this paper showing mounting positions of all the sensors (right).....	173
Figure 77. Structured Light scan volume on lunar test terrain.....	173
Figure 78. Linear gray code scanning of the lunar terrain causes errors at the finest resolution due to differences in underlying scene albedos (top). Image cues can direct the scanner to resample salient regions such as rocks a pixel at a time (bottom). Samples are shown simultaneously illuminated for illustrative purposes; scanning occurs one dot at a time.....	174
Figure 79. Many images of the scene taken from a single view point with moving sun can further enhance the super-resolution approach by utilizing photometric stereo. (1) A low resolution LIDAR model is fused with the shading estimates from multiple images using photometric stereo (2) to create a super-resolution model.....	177
Figure 80. Normal maps of the moonyard estimated from intensity images. Estimating normals from many images using photometric stereo produces cleaner, more accurate normals than the single image technique. Moreover, seeding the tilt angles with sparse range information is no longer necessary.....	179
Figure 81. Shadow Maps can clearly identify shadows and occlusion edges (left). Detected edges (right) can be expanded to create a cost map in MRF fusion to discount for inaccurate interpolation data. Note that the color images shown are not the shadowed images used. These ambient images are used for visual clarity. The position of the sun goes from left to right.....	180
Figure 82. Materials can be identified with physics-based segmentation techniques. The Marius Hills skylight (left) has anomalous features that can inform the geology of the region. The image on the (right) shows a rock from the moonyard where regolith has brushed off due to the steep slopes. Lumenhancement can identify that it is a different material than the background, while not being fooled by other regolith covered rocks – a task difficult even for a human.....	182
Figure 83. Simulated Rendering of Moon Terrain. Materials from this thesis are utilized to generate fidelity of appearance. The underlying regolith-covered terrain uses a Torrance-Sparrow model,	

while the rocks use a Lambertian model. The physical model is joint work documented in [Jones, et al. 2012]. 183

Figure 84. Simulated Orbital Imagery of the Moon from Two Sun Angles. The top row simulates the effect of pixel noise in the camera while the bottom images are ideal. Contrast is enhanced here for viewing. Even then, it is almost impossible to distinguish rocks in the images because they are small and have similar appearance to the regolith. Hint: red arrows point to some rocks which appear as bright blips. 185

Figure 85. Comparison of Segmentation between a Naïve and the proposed Material Approach. Naïve segmentation uses an LAB transformation and thresholding, while the material approach uses photometric stereo from several images. Both algorithms use mean shift clustering. 186

Figure 86. Number of Rocks Detected vs. Accuracy of the Cluster compared to the true region. 188

Figure 87. Super-resolution modeling of this mine wall demonstrates a generational leap of model quality using Lumenhancement. 189

Figure 88. Reference Material BRDF Polar Plots. 203

Figure 89. Underground Material BRDF Polar Plots. 204

Figure 90. Planetary Material BRDF Polar Plots. 204

Glossary of Terms

albedo – A ratio of reflected light to incident light that is an intrinsic property of materials.

appearance – Light interaction with a scene that can be captured with imaging sensors.

brdf – The Bi-directional Reflectance Distribution Function, which encodes reflection of light from surfaces. The BRDF is the principal unique property of materials in computer vision.

color – Spectral variation of albedo across wavelengths of light. In graphics, it is a vector of tristimulus (RGB) values when used in context with an observer and known spectral response functions.

diffuse – Scattering radiation equally in all viewing directions.

emergence – Outgoing light ray from a surface.

incidence – Incoming light ray to a surface.

inversion – Use of the BRDF and radiance observations to recover the intrinsic properties of a body, i.e. parameter fitting.

isotropic – (1) Radiating equally in all directions. (2) Rotationally invariant in the context of BRDFs. Opposite of **anisotropic**.

model – A three dimensional, geo-registered map of the environment. Data is not limited to just geometry.

macroscopic – Robot or human scale; *macroscopic features* are differentiable with current surface sensing technologies.

microfacet – Approximation of flat surfaces as patches consisting of flat microscopic cavities.

radiance – Light leaving a surface. Opposite of **irradiance**.

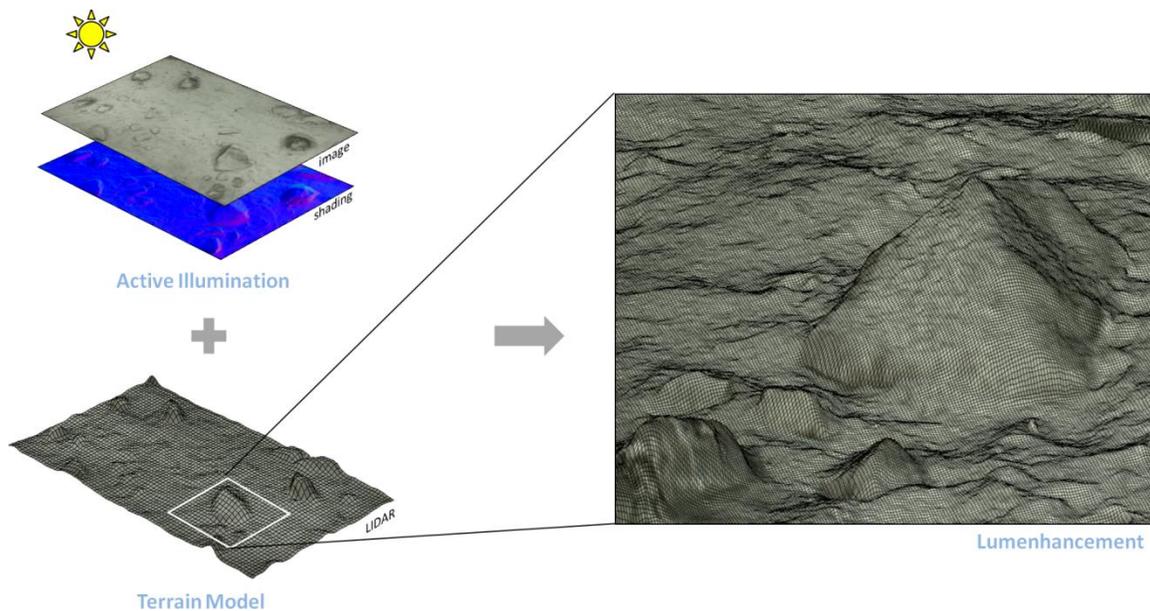
perceptual – Relating to sensed phenomena as opposed to physical.

specular – Glossy; A surface with a peak radiance in a viewing direction at opposition to the light source, resulting in “highlights”.

variegation – consisting of many discrete materials or colors, “splotchy”.

Chapter 1:

Introduction to Lumenhancement



Robotic mapmaking is destined to impact the safety, economy and science of operations on earth and other worlds. Maps are geospatial representations of geometric, physical and visual information. Robots use maps to document, reason, and interact with the world; and maps are created as a byproduct of these actions. Unlike the two dimensional maps of yesteryear and the 2.5D occupancy grids of today, the future demands dense three dimensional geometry geo-registered with multi-spectral and multi-sensor information. These are not maps in the traditional sense, but comprehensive *models*: they provide environmental understanding beyond simply geometry.

Only robots are able to generate models with superb level of detail and quality, and to do so in hazardous environments that humans cannot visit. In fact, modeling is *the* primary motivation and purpose for robots tasked with planetary exploration, survey and inspection. Such applications represent the prime unexploited opportunity for robotic mappers; their prevalence throughout the universe is astounding. However, planetary environments are also distinguished by critical constraints on power, mass, surface reflectivity, and ambient illumination available for sensing. Improvements in modeling efficiency and quality will represent significant opportunity for robotic technology in this domain.

Lumenhancement, a concept introduced in this thesis, can be utilized in the modeling process to achieve these objectives. Lumenhancement is simultaneously an approach and a philosophy. It is an exploitation of optical domain knowledge with targeted illumination and also a data-centric view in the use of robots as modeling tools. These two concepts are inextricably linked by the laws of image formation as they govern the appearance of environments. As a result of this physical grounding, Lumenhancement is both effective and broadly applicable as experimentally demonstrated herein.

1.1 The Case for Planetary Environments

This thesis targets modeling in the planetary domain. In this context, *planetary* encompasses both barren cosmic landscapes and terrestrial underground voids. Quantitative evidence is presented in this document as to why these seemingly disparate environments are interchangeable and equivalent. Moreover, this determination will have unique ramifications in the approach to modeling. However, at this point it is sufficient to simply convey intuition that these environments are optically congruent in addition to being similarly remote, hazardous, and map relevant. Furthermore, this duality will present an interesting breadth of applications to demonstrate efficacy where possibility for experimentation may be limited - sending a robot to the moon for example.



Figure 1. Robotic Modeling in Planetary Environments. Examples include (1) automated lunar landing, (2) terrestrial underground inspection, (3) lunar resource extraction, (4) mapping of “skylights”, (5) exploration of abandoned mines. [Photo courtesy Red Whittaker, CMU for (1); David Wettergreen, CMU for (2); all others Uland Wong].

Planetary environments are among the most hazardous, remote and unexplored in the solar system. Yet, they are also among the most likely candidates for harboring extraterrestrial life [Thompson 2008], finding volatiles [Wettergreen, et al. 2009], and providing warmth and radiation shielding for human explorers [De Angelis, et al 2002]. On earth, the underground presents tremendous material and scientific value as well as ecological danger and security risk [Morris 2007], [Omohundro 2007]. There is urgent need to explore, document, and evaluate planetary spaces with robots and to do so in a superior and economic manner beyond the state-of-the-art.

Much of prior work has focused in adapting general surface sensing techniques to extreme environments. Very little research has been undertaken to address the effectiveness of these methods in the domain and fewer still have addressed notions of optimality or optimization of sensing resources. These issues are particularly important because **many of these environments where robots are prime candidates for mapping are also those that are acutely difficult for sensing and robot operation**. Many planetary bodies, like the Moon, asteroids, and Mercury exist in the extremes of illumination. Particulates in submerged or dusty atmospheric environments scatter light sources and confound both

image-based and time-of-flight measurement techniques. Lastly, lack of bandwidth, power and mission lifetimes, limit the type and quality of sensing available.

Fortunately, planetary environments are not without advantages which can be leveraged to maximize modeling productivity. Uniform, barren, rocky surfaces are known to be uniquely suited for use in image-based perception methods. While natural illumination fluctuates between the harsh limits of absolute dark and unadulterated intensity, it is in fact physically simple and mathematically factorable for many uses. Physical constraints on the distributions of both the macroscopic and surface geometry lend themselves to characterization and utilization in prediction of performance. The close relationship between planetary spaces to many outdoor domains could ensure wide applicability to robotic mapping in general. These reasons, among others are compelling motivation for the research and development of planetary-specific sensing.

This thesis seeks to answer the question:

How can the appearance of planetary environments be exploited to improve geometric modeling in a general manner?

Research Question

Specifically, this thesis uses *exploited* to mean “in the sensing process” – independent of autonomy or robot function and utilized in a strap-down manner. *Generality* is a consideration of broad applicability for geometric modeling and the expected favorable performance of these techniques in many related environments as opposed to single mission use.

1.2 Shortfalls in Planetary Modeling

Planetary applications have produced several notable victories for orbital mapping. The Mars Global Surveyor has generated complete altimetry on a planet-wide scale [Albee, et al. 2001]. The moon has been mapped with a combination of imagery and altimetry from the Lunar Reconnaissance Orbiter [Smith, et al. 2010] and half-century old Apollo data using modern long-baseline stereo techniques [Edwards, et al. 2006]. Orbital probes have also visited Mercury [Hawkins, et al. 2007] and asteroids like 433 Eros [Zuber, et al. 2000]. The work of Kirk in generating high resolution topography of Mars using

HiRISE [Kirk, et al. 2007] and establishing the lunar control network [Archinal, et al. 2006] is a major step toward earth-like mapping of planets. However, these approaches are limited by comparatively low resolution, single-perspective¹, and remote observation which are unsuited to human-scale activities.

The only significant examples of robotic planetary modelers are the Mars Exploration Rovers, which mapped traces of the red planet in high detail over long duration using stereo vision and photogrammetry [Goldberg, et al 2002], [Cheng, et al, 2006]. Here on earth, robots are not yet socially and economically accepted for subsurface use even in the most dangerous of applications. However, inroads for modeling have been created through the pioneering research of [Champeny-Bares, et al. 1991], [Omohundro 2007], [Morris 2007] for mining and [Johnson, et al. 1996], [Yoerger, et al. 1999], [Fairfield, et al. 2005] for underwater application.

While there is no doubt more missions will target these planetary spaces for years to come, the current state of robotic planetary modeling stands in stark contrast to the proliferation of terrestrial surface and indoor mapping. Robotic modeling technology is omnipresent in safe driving applications like the DARPA Grand Challenges [Thrun, et al. 2006], [Urmson, et al. 2007]; large-scale human demography like Google Street-view [Anguelov, et al. 2010] and Google Earth [Kennedy 2009]; building survey [Oliver, et al 2011]; indoor [Tardos, et al. 2002] and social robotics [Izadi, et al. 2011]; and even micro-aerial inspection [Chambers, et al. 2011]. These maps are expansive, highly accurate, and easily updated because their creation represents simple sensor application in highly structured environments. Copious computing, power and sensing resources are easily brought to bear, without regard for efficiency.

This approach cannot succeed in many field environments – let alone planetary - where impediments to sensing are ubiquitous. Environments which are fundamentally different optically, geometrically and physically, demand fundamentally different approaches. Amazingly, terrestrial surface sensing techniques and commercial sensors tailored to industrial application are commonplace for planetary operations. The truth is that techniques for *optical sensing* - the primary means of model making - have lagged behind the other capabilities of robots in this domain. Terrestrial approaches might work or even suffice for many applications, but they are generally doomed to ineffectiveness.

The costs of model acquisition are drastically disproportionate to the sum of knowledge gained. This knowledge gap is the result of unprincipled measurement, inflexible sensors and poor visualization compounded by the difficulties of the domain. Robot maps vary widely in effectiveness according to

¹ Orbital approaches are 2.5D and grid based

application: volumetric estimation depends on terrain structure and sensor viewpoint; measurement scale significantly affects feature identification; and virtually no attempts have been made to improve human readability of models beyond raw point displays or meshing. These considerations, among others, have particularly contributed to the slow adoption of robotic mapping technology underground, for example. Technologies which improve the real and perceived effectiveness of robot mapping will reduce the stigma of “expensive toys” and potentially save lives in the process.

1.3 Document Overview

This thesis introduces the idea of Lumenhancement to the sensing literature. The document comprises seven chapters: (1) Introduction, (2) An Optical Interpretation of Modeling, (3) Thesis Statement, (4) Related Work, (5) A Characterization of Planetary Appearance, (6) Case Studies for Lumenhancement and (7) Conclusion. The introduction – this section – describes the technical gap in planetary modeling and motivations for using Lumenhancement to improve robotic sensing in the domain. Chapter two reviews image formation fundamentals and establishes the concept of appearance spaces, which are a vehicle for analyzing the applicability of vision techniques. Chapter three is the thesis statement which defines Lumenhancement and introduces the three topical areas that are foundational to the thesis: domain knowledge, optical sensor fusion and active illumination. Chapter four describes related work in each of the topical areas, explains foundational concepts, and contrasts this thesis with prior work. Chapter five documents the experimental characterization of the planetary domain for the parameters of appearance using gonioreflectometry and range sensor survey. Case studies, chapter six, presents three example methods for improving modeling with Lumenhancement and provides experimental evidence for the efficacy of these techniques. Lastly, the conclusion covers the contributions of the thesis and its significance to planetary modeling.

Chapter 2:

Foundations of Appearance in Modeling

“There is the motion, the actual wave and radiation of the darted beam; not the dull universal daylight, which falls on the landscape without life, or direction, or speculation, equal on all things and dead on all things; but the breathing, animated, exulting light, which feels, and receives, and rejoices, and acts; which chooses one thing and rejects another; which seeks, and finds and loses again, leaping from rock to rock, from leaf to leaf, from wave to wave, glowing, or flashing or scintillating, according to what it strikes.”

John Ruskin, *Modern Painters*, 1843

on the nature of *chiaroscuro* - the use of contrasting light to emphasize geometry in painting

2.1 Optical Sensing in Modeling

The classical masters of the Renaissance and Baroque knew the power of controlled lighting in defining boundaries, conveying volume, and elucidating features. Despite the incredible complexity of physical light interaction and projection from three dimensions onto a two dimensional image plane, the

information – and emotional – content of *chiaroscuro*² paintings could not be more clear. Similarly, a light detector provides the ability to surmise shape, substance, and spatial information from a distance, at high information density, and without alteration of the scene. It is not a surprise then, that robot modeling of this generation³ is almost exclusively optical. In fact, this is also true in the vast majority of robotic exteroception.

The dominant population of optical sensors can be classified into two types: range sensors and image sensors. Almost any number and combination of these classes can be found on modeling robots - Figure 3 illustrates some optical configuration examples. Range sensors like LIDAR, RADAR and Structured Light measure distance to objects, most often by generating and measuring controlled illumination (see background section for an in-depth review). This distance reading is coupled with the known pose of the sensor to produce a single point sample in three dimensions. While their mode of operation is often beam based and areal, samples are usually treated as infinitesimal Dirac “points”. With enough samples, a representation of the scene surface geometry emerges either through connective (i.e. meshes) or volumetric (i.e. clouds, voxels) means. Range sensors are the primary means of robotic mapmaking because they measure geometry directly and at the heart of maps are a collection of geometric relationships between objects.

² Chiaroscuro, meaning “light and dark” in Italian, is a form of painting which emphasizes contrasting illumination for dramatic value.

³ Acoustic sensors have only remained relevant in underwater modeling and hobbyist robotics. Tactile sensing is invasive and irrelevant for rapid modeling in most field environments.

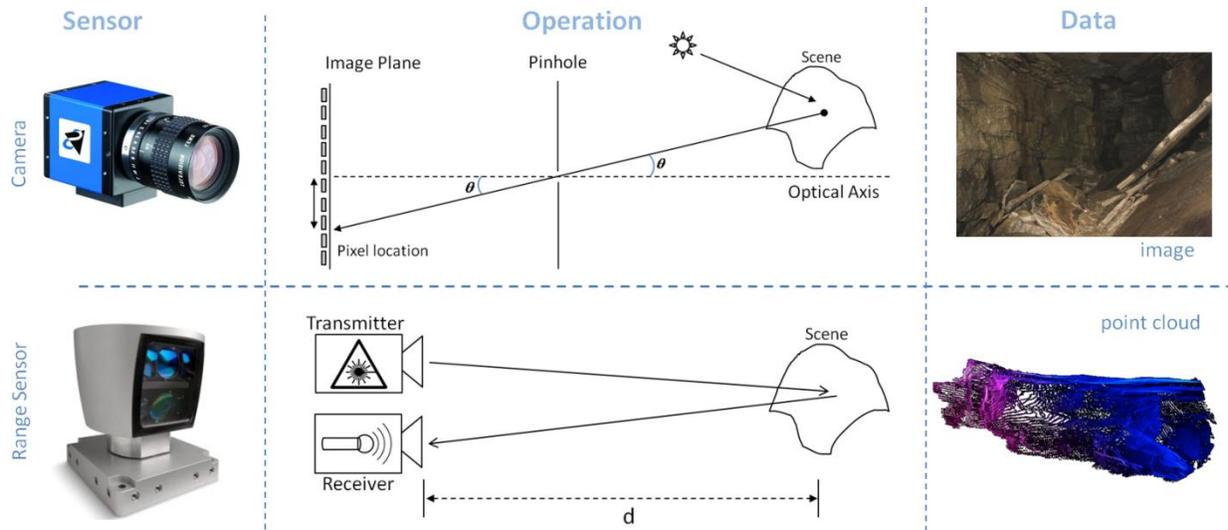


Figure 2. Canonical Optical Sensors. Cameras measure a composition of material, geometry, and illumination using the mechanism of two dimensional image projection (top). Only angle and intensity information for each light ray is preserved. Range sensors directly measure geometry to produce “clouds” of point samples in three dimensions (bottom). Data is from a sandstone cave.

Cameras, which are image sensors, capture the appearance of a scene as a two dimensional projection of light intensity. This appearance is a composition of light phenomena, geometry, material and color spectra. Traditional cameras are passive sensors in that they measure only light reflected off the scene and have no indication of the original timing, frequency, or intensity. Such environmental illumination can be from the sun, another natural source or an artificial source. Light used in this manner provides the possibility of confusion; its effectiveness as a sensory medium is based on consistent probabilistic correctness. Ruskin’s observation of the ineffectiveness of general daylight as compared to more discriminating directional illumination summarizes one facet of this problem. As the compositional and projective processes of image formation are lossy in nature, extraction of meaningful information requires the use of assumptions and the validity of these assumptions [Horaud, et al. 1988]⁴.

Despite these drawbacks, cameras are the only commodity⁵ sensor capable of detecting material or color – even if these properties cannot be unambiguously discriminated from an image. Camera data closely matches the human visual experience. The measurement density and throughput of cameras is also far superior to range sensing as a result of prevailing physical designs. Images are often used in

⁴ The reader is encouraged to consult a sensor physics reference like [Fraden, 2003], for more detailed explanation of camera and range sensor theories of operation.

⁵ Many types of spectrometers exist for detecting material properties. While highly accurate in classification, these are not conducive to mobile 3D mapping due to low measurement density and invasive sampling.

detailed understanding of the scene for these reasons. However, any geometric information gleaned from images would greatly supplement the low density, slow acquisition of range sensing.

This work blurs the distinction between range sensors and cameras by exploring beyond the traditional uses of images in mapping. In particular, superior recovery of geometric and appearance information is demonstrated from images taken under controlled illumination and constrained appearance properties.

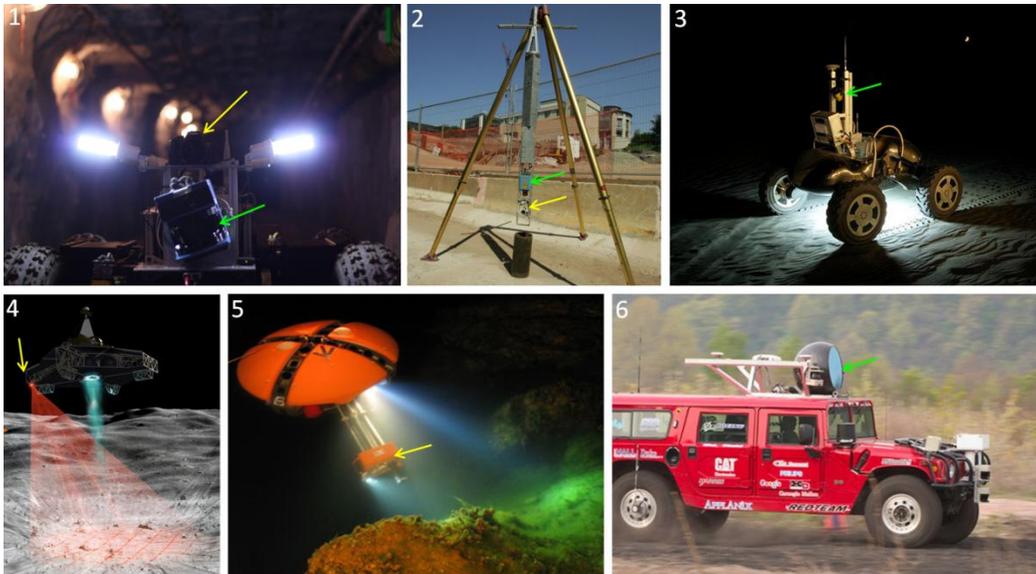


Figure 3. Modeling Robots and Their Sensors. Some examples of field robots utilizing LIDAR ranging (yellow arrows) and cameras (green arrow) in varying configurations. (1) Cavecrawler, a mobile mine inspection robot; (2) Ferret, a snake-like borehole inspection robot; (3) Scarab, a lunar prospecting robot [photo courtesy David Wettergreen, CMU]; (4) an automated lunar lander concept; (5) Depth-X, an underwater explorer that utilizes sonar and a camera [photo courtesy David Wettergreen, CMU]; (6) H1ghlander, an off-road racing robot [photo courtesy Red Whittaker, CMU].

2.2 Domains as Appearance Spaces

A colloquial definition of the word *domain* has thus far been used in this thesis and prior work to denote a province of robotics. In this sense, domains are collections of exemplary environments, associated by their similarity of application. An “indoor” domain, for example, may comprise homes, office spaces, and restaurants, where an intersection of applications includes housekeeping and people interaction. This thesis augments prior definitions with a novel consideration of environments which are similar in appearance. These *optical domains* are central to reasoning about planetary sensing approaches and their limits of applicability.

The groundbreaking work of [Omohundro 2007] and [Morris, et al. 2006; 2007], were some of first to define domains in field robotics and to utilize their properties in development of robots. Their work serves as specific inspiration for this thesis, as they first described a subterranean domain, made compelling arguments for traveling underground and initiated a campaign for autonomous underground mapping. Omohundro, in particular, was the first to attempt formal characterization of underground spaces beyond qualitative arguments. In his thesis, he explores a spectrum of range sensing methods and classifies underground voids by immersive material: water, air, collapsed or solid. Internal robotic exploration and modeling is distinctively identified as a more effective alternative to remote surface sensing methods like ground penetrating radar (GPR). Based on these observations, a mechanical configuration taxonomy sorted by ingress method and portal size is proposed for robots designed to map the domain. The term “subterranean domain” is coined and implied to mean terrestrial mines, tunnels and caves, with example implementations primarily targeting application in abandoned mines.

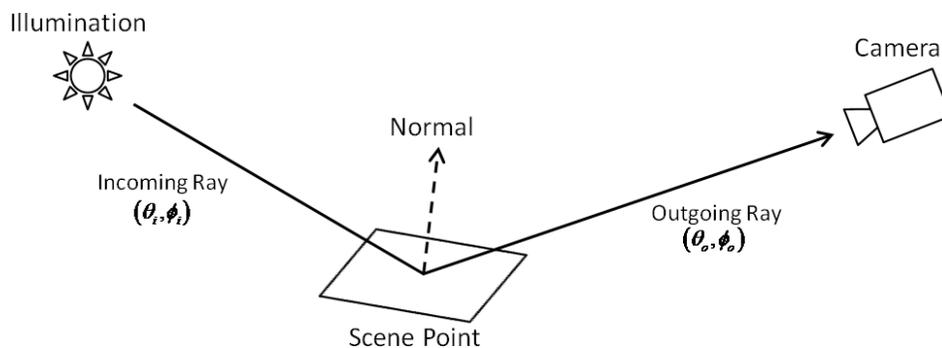


Figure 4. Surface Appearance Model. Incoming light rays, parameterized by polar angles, strike the object surface. These rays are decomposed and reflected in many directions determined by the material properties. Some of the outgoing rays will reach the sensor location indicated by the viewing direction.

Though the primary motivation of Omohundro’s work is sensing of air-filled underground voids, the assessment and characterization of environments presented lacks a sensing context. When the geometric nature of mines (including ingress type) is considered along with the material and atmospheric characteristics, a more general but nuanced picture of the subterranean domain is revealed. Many mapping methodologies that apply to abandoned mines and caves also apply to a plethora of other environments, both on earth and other worlds. Yet, between two dry mines on earth the same LIDAR range finder may work well in a limestone mine, but fail to return readings in a coal mine where anthracite coal exhibits strong absorption of near-infrared signals for example.

Optical sensors like LIDAR and cameras are central to robotic mapping in almost every environment. As such, the parameters of *appearance* provide a much better basis for understanding and describing domains, particularly, when the focus is mapping. Geometric optics - rules that approximate light propagation - governs surface appearance through the principles of reflection and refraction. In the classic appearance model (Figure 4), light is radiated from an illumination source along geometric rays. These light rays intersect scene points where they may be partially absorbed and reflected in many directions (scattering). The mechanism of reflection is a complex function of the macroscopic geometry as described by a surface normal and the optical properties of the material which define the intensity and distribution of scattered light. Light from the scene may further reflect multiple times before passing through a lens, where the rays are warped through refraction and are finally captured by a sensor [Horn 1986]. This simple geometric model does not have the expressive power to describe phenomena such as diffraction or participatory media, though the former is rarely a major issue in imaging and latter can often be approximated with greater light source complexity in scenes where sensing distance is comparatively short⁶. For the express purpose of describing application domains, however, the aforementioned approximation suffices.

Some of these parameters can be manipulated or controlled in the design of an optical perception system, but others must be accepted as immutable. It is evident that natural illumination and the parameters of reflection, scene geometry and material, are functions of the environment, while refraction (lensing) and detection are properties of sensor construction. This dichotomy forms the basis for defining optical domains in this thesis; environments in the traditional sense are described by the circumstances of their appearance - parameterized by geometry, materials, natural illumination and participatory media. Domains are not just collections of characteristic environments, but a continuous “appearance space” spanned by these constituents.

⁶ Daylight illumination is often approximated as a combination of directional lighting from the sun and ambient blue light from the sky for the purposes of vision and rendering. The effect of atmospheric Rayleigh scattering is minimal at robot sensing distances; however larger particles such as dust or smoke in the Mie domain can be detrimental to sensing.

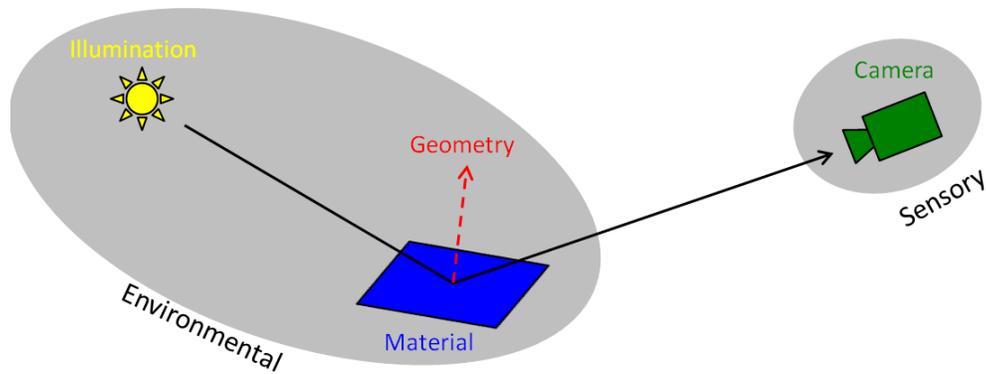


Figure 5. Parameters of Image Formation can be divided into two classes: those intrinsic to the application environment (material properties, surface geometry and natural illumination) and those which are byproduct of sensor design.

In this framework, environments are only differentiable with respect to their optical effects in image formation. Conversely, environments with optical similarity within the tolerance of the sensing resolution or their respective natural variations are indistinguishable. This judgment is regardless of traditional factors such as geographic location or ingress constraint because those are not detected by sensing. This classification creates a convenient way of analyzing the generality of sensing techniques as the applicability to environments directly corresponds to the space of optical phenomena spanned.

Consider a graphical representation of the appearance space, where three arbitrary orthogonal axes represent complexity of geometry (x-axis), material (y-axis), and natural illumination (z-axis). Increasing distance from the origin denotes increasing complexity of each attribute. The natural illumination axis describes the potential of sunlight (or derivatives like airlight) to interfere with sensors. Simple illumination environments include those where the sun can be modeled as a point source, or where no natural light exists (i.e. caves), while complex environments exhibit multiple scattering and atmospheric effects (i.e. terrestrial outdoor daylight scenes). The material axis describes the quantity of materials – and consequently reflectance functions – that occur with significant frequency in the environment. Modern coal mines, with metallic roof beams, nylon curtains and mine equipment are material-complex, while asteroids, generally mono-material and covered in regolith, are simple. Lastly, geometry axis measures both the minutia of surface features as well as the quantity of concavities and convexities encountered in the structure. Underwater voids are quite geometrically complex due to formation from flowing liquid, while lava tubes, are not much more than smooth, lazily wandering cylinders.



Figure 6. Mondrian Environments describe the simplest conditions of image formation which are matte materials, ambient illumination and planar surfaces.

The simplest conditions of image formation – matte materials, ambient illumination, and planar geometric surfaces – meet at the origin of the appearance space. These conditions are exemplified by Mondrian environments [Blake 1985], named for the Dutch Neoplastic artist Piet Mondrian, who painted only black-lined grids with variably positioned red, yellow and blue rectangles. While Mondrian environments are an idealization, they are not unlike indoor spaces which are frequent targets of camera-based sensing. These minimalist approaches are often successful as a result of this environmental simplicity.

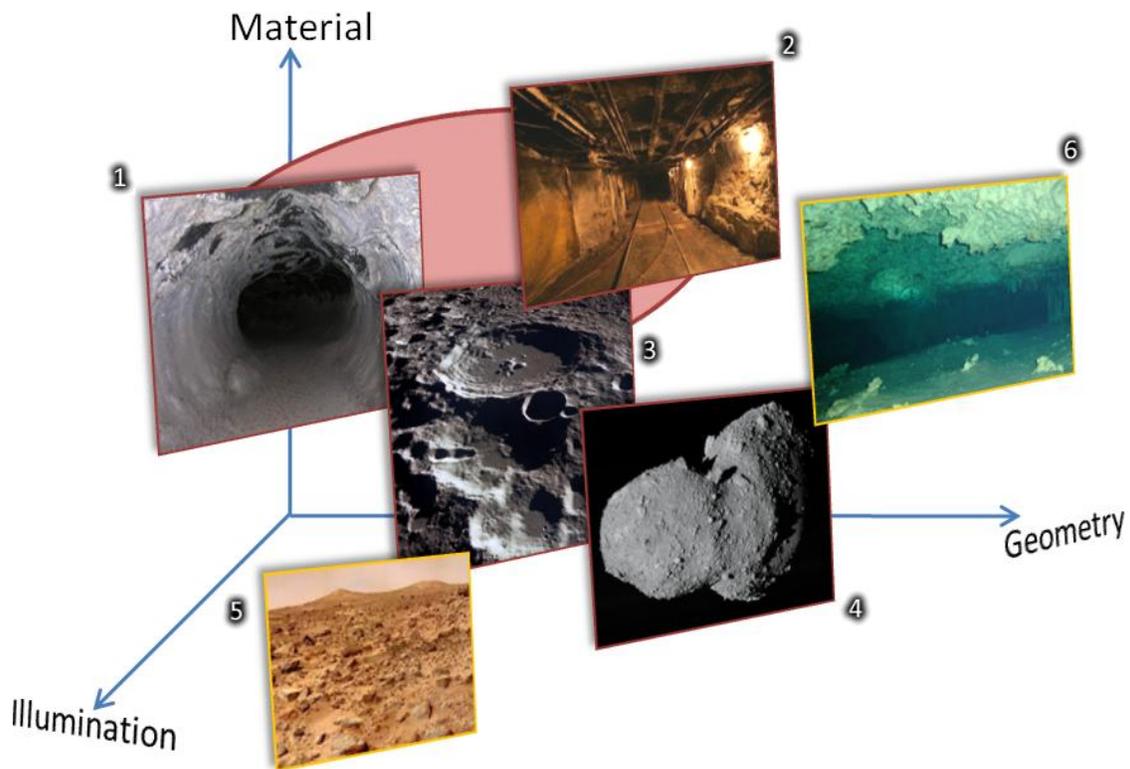


Figure 7. The Planetary Domain and Related Environments. The axes denote increasing quantity and complexity of materials; geometry, which includes macro-scale features and minute surface detail; and illumination – number and complexity of natural light sources. The origin represents a perfect Mondrian environment. Other environments shown include: (1) Lunar and terrestrial lava tubes, (2) underground mines, (3) Lunar craters, (4) asteroids, (5) the Martian surface, (6) the sea floor and underwater caves. Work presented in this thesis applies to (1-4 in red), while the subterranean domain of Omohundro and Morris is illustrated with the red oval.

The subterranean domain of Omohundro might occupy a region near the origin of the illumination and geometry axes, while being moderately elevated in the material axis (z). Figure 7 shows subterranean environments as an abstract red oval in the appearance space along with photographs of some characteristic environments. A lava tube and a coal mine (shown) are two examples that fall under the original strict definition of this domain. However, it is quickly apparent that some “surface” environments on asteroids or the moon and partial voids such as lunar craters or skylights⁷ are at least as optically similar to both these examples as the examples are to each other. In contrast, other terrestrial underground environments like flooded caves are less similar in appearance.

⁷ Lunar Skylights are recently discovered features thought to be openings to lava tubes and formed through surface collapse [Haruyama, et al. 2009]. It is believed the openings at the top of the void are much smaller in diameter than the void itself, presenting an interesting “skylight” illumination scenario. It is unknown whether the rest of the lava tube is accessible from these locations.

This thesis defines the *planetary domain* to be the set of barren, dry, rocky environments with simple natural illumination. These qualifiers and their significance to appearance are now discussed:

Barren. Planetary environments have a few dominant surface materials which can be approximated as smooth and locally planar. These materials are spatially clustered and discriminable using vision techniques. Complex man-made artifacts and materials are virtually nonexistent. This property means the environment is conducive to measurement and that reasonable inferences can be made about data which is interpolated between sensor readings or within the integrating cone of a single reading. Assumptions that small changes in perceived intensity are the result of shading while abrupt changes are due to material or object boundaries are valid.

Dry. There is no participatory media, like dust or smoke, in the environment which can appreciably absorb or scatter illumination between the scene and sensor. This requirement is relative to robot sensing distance and scale. Air-filled voids are only minimally different from vacuums for most sensing purposes, but water or smoke filled voids are not. This property ensures that illumination used or carried by the robot reaches the scene with the same distribution that it was generated so that this information can be utilized for vision. Furthermore, the signal from the scene is not unduly corrupted by the participatory media.

Rocky. Surface materials are mostly diffuse such that robots are not required to reason about complex phenomena like translucency, mirror specularities and light emission. Variegation due to mineralogy or aggregation is macroscopically approximable as a single *albedo* or color. This constraint enforces strong correlation between image features and shape and enables simple geometric recovery given image data. Moreover, diffuse reflection ensures that scene features remain highly correlated between similar, but different viewpoints.

Simple Illumination. Natural illumination can be approximated as a linear combination of point sources of known location or is absolutely dark such that an artificial source of these properties may be carried by the robot. Any ambient or stray light in the scene must either be sufficiently minimal or removable; this may summarize the “void” attribute of Omohundro and its benefit to sensing⁸. The view of this thesis is that illumination may be environmental, but is not an explicit

⁸ Bounded, concave voids result in a constant ambient illumination from interreflection when the void is small, or negligible interreflection when large.

property of scenes where modeling is concerned. Thus, it is one of the few factors that can be utilized to control image formation such that these properties can be discriminated. Simple Illumination might also be called “dark” for a less precise, but helpful single-word mnemonic qualifier.

This domain encapsulates the original planetary applications of interest and many of the underground environments considered by Omohundro with terminology that is context neutral in regard physical location or configuration. While this redefinition may seem like an exercise in semantics, it is critical to conveying that optical sensing techniques may have significance beyond their original application intent. In particular, this work rejects that traditional “void sensing” techniques must only be limited to classes of enclosed environments like a mine. Instead, it is more important that geometry (like material and illumination) is consistent and characterizable. Indeed, skylights and craters are only trivially different from voids, as all three are macroscopically flat.

With these criteria, it is possible to observe an intra-class variance in the planetary domain and also a continuum of similarities between it and related environments. The boundaries of this domain are approximate and ultimately insignificant; in quantitative analysis, a distance measure is arguably more useful than arbitrary thresholds. Example subterranean environments include coal mines, lava tubes, lunar craters, lunar skylights, and asteroid surfaces. Near neighbors include the Martian surface with more complex, atmosphere-scattered solar illumination and flooded voids with no natural illumination, but strong participatory media and non-planar geometry. Planetary modeling strategies may pertain to these neighbors with some modifications or decreased effectiveness. Table 1 below summarizes several characteristic environments as well as the neighbors shown in Figure 7. These determinations are not rigid and are intended only to serve as an informative guide. For example, some polar explorer concepts on the moon circumnavigate at latitude to produce maximum solar energy [Wettergreen, et al. 2005]. Such a robot would always shadow the terrain in front of it, and would be operationally identical to actively illuminated crater explorers.

Table 1. Summary of Characteristic Environments and Neighbors. Environments directly applicable to this thesis are in highlighted in blue. Although the underwater realm and Martian surface are distinct from the other environments, approaches addressed here are possibly applicable.

Environment	Materials	Geometry	Natural Illumination
Coal mine	Moderate, a few dominant, covered in dust	Moderate, predictable	None
Terrestrial cave	Moderate, mostly diffuse	Moderate, bounded	None
Lava tube	Few	Simple, smooth and straight	None
Lunar crater	Few, regolith and bedrock	Moderate, planar	Simple, starfield and sun, some permanently shadowed
Lunar skylight	Believed to be few	Moderate-Complex, tunnel-like	Simple, bounded starfield and sun, may be shadowed
Martian surface	Moderate, dust and dominant rock types	Moderate, planar	Moderate to complex, depends on existence of atmosphere/weather
Underwater	Possibly many	Complex, planar	Complex, immersed scattering media, no natural light at depth

2.3 The Role of Domains

Assumptions are crucial elements in simplifying robotics perception problems. Line-based features, which are uncommon in the natural world, are omnipresent in indoor feature tracking applications. In effect, there is an (often true) belief that only geometric changes result in a change of intensity. Many surface robots represent the world as 2.5D grids because they likely do not leave the ground plane. Stereo vision approaches assume changes in viewpoint from parallax do not change incident illumination. Perception cannot occur without requiring some set of assumptions to be mostly valid during the majority of the application lifetime.

While planetary environments are filled with challenges, assumptions can also be made to enhance sensing. However, the set of suitable assumptions is fundamentally different from that made in indoor or surface environments. The domain classification exercise provides a vehicle for constructing assumptions and analyzing their validity. An understanding of optical properties can be exploited to

enhance perception with targeted sensing. Consider a numerical formulation of the geometric optics given by Horn in section 2.2, Figure 4. The radiance $L_o(x, w_o)$ at position (x) , with incoming and outgoing vectors (w_i, w_o) , and surface normal (\hat{n}) is given by the Rendering Equation [Kajiya 1986]:

$$L_o(x, w_o) = L_e(x, w_o) + \int_{\Omega} f(x, w_o, w_i) L_i(x, w_i) (w_i \cdot \hat{n}) dw_i \quad (2.1)$$

The Rendering Equation describes light transport for geometric optics.

where, $f(x, w_o, w_i)$ is the material reflection function, $L_i(x, w_i)$ is the illumination distribution, and $(w_i \cdot \hat{n})$ is the projected irradiance onto the surface. $L_e(x, w_o)$ is the light emitted from the scene (i.e. the scene is a light source), which is not further considered in this work. It would be beneficial if this equation could be reduced to a simpler form, such as a triple product:

$$E = M \times I \times G \quad (2.2)$$

where E is the radiance as perceived by a sensor, $M = f(x, w_o, w_i)$ is the material, $I = L_i(x, w_i)$ is the illumination and $G = (w_i \cdot \hat{n})$ is the geometry. Marginalization can occur with knowledge of any factors to enable probabilistic estimation of unknowns. With sufficient knowledge, certain critical constraints, or multiple observations, even unambiguous recovery is possible. Noiseless decomposition of individual terms is not likely, however, as assumptions and constraints are themselves subject to error.

This form of qualitative reasoning can be applied to planetary environments to recover geometry given the aforementioned assumptions of **barren, dry, rocky**, and **simple illumination**, for example. Consider the rendering equation (2.1) once more. Applying assumption of **simple illumination** (external, point sources and direct illumination), the equation becomes⁹:

$$E(x) = L_o(x, w_o) = L_e(x, w_o) + \sum_k f(x, w_{o,k}, w_{i,k}) L_{i,k}(x, w_{i,k}) (w_{i,k} \cdot \hat{n}) \quad (2.3)$$

⁹ This Interreflection of light within the scene is assumed to be negligible compared to the direct component.

A hemispherical integral is no longer required. Instead, the contribution of a discrete number of k infinitesimal point sources can be counted and summed linearly. Moreover, light from the scene is either emitted to the infinite void or reaches the camera directly such that $L_o(x, w_o)$ is equivalent to $E(x)$ - the intensity image of perceived radiance in the corresponding set of pixel directions w_o . The rocky surfaces are also assumed to be non-emitting, meaning that only directly reflected light from a source not in the scene reaches the sensor, giving:

$$E(x) = \sum_k f(x, w_{o,k}, w_{i,k}) L_{i,k}(x, w_{i,k}) (w_{i,k} \cdot \hat{n}) \quad (2.4)$$

Rocky surfaces are also assumed to be diffuse. Thus the reflectance function is replaced with a constant term $\frac{\rho}{\pi}$, which does not change between any incoming or outgoing angles. The significance of this scalar Lambertian albedo and its normalizer is described in later chapters.

$$E(x) = \sum_k \frac{\rho}{\pi} L_{i,k}(x, w_{i,k}) (w_{i,k} \cdot \hat{n}) \quad (2.5)$$

Next, planetary environments are assumed be **dry**. There is no participatory media which can affect the angular distribution or intensity of light reaching the scene. Thus, the light source function can simply be replaced with the scalar intensity of each point source:

$$E(x) = \sum_k \frac{\rho}{\pi} I_k (w_{i,k} \cdot \hat{n}) \quad (2.6)$$

Lastly, applying the **barren** assumption, constrains the surface geometry to be local smooth and hence differentiable. Given a surface parameterization of $f(x, y) = z$ and its partial derivatives (f_x, f_y) the equation becomes:

$$E(x) = \sum_k \frac{\rho}{\pi} I_k (w_{i,k} \cdot \nabla x) \quad (2.7)$$

The surface normals, which control incident illumination and consequently shading, are related to the macroscopic geometry only through the gradient operator¹⁰. Lastly, if a single point source, such as the sun¹¹, is assumed, the equation reduces to:

$$E(x) = \frac{\rho}{\pi} I(w_i \cdot \nabla x) \quad (2.8)$$

$$E \propto \cos \theta_x \quad (2.9)$$

Thus, the perceived image intensity (E) is variable with respect only to the surface gradient of geometry (∇x) and the cosine projected angle of the source. The complex integral rendering equation has reduced to the desirable form of equation (2.2), if the planetary assumptions are valid. These constraints are almost perfectly suited for recovering surface (normal) geometry from intensity images. This shape-from-shading framework is among the simplest of vision based algorithms, yet there is no record of utilization in these environments. Perhaps the lack of a pragmatic solution to dealing with possible albedo changes ($\rho \Rightarrow \rho(x)$) and boundary conditions ($\nabla x \neq \hat{n} \quad \forall x$) have precluded prior attempts. A modification of this algorithm which handles these conditions is one of the techniques presented in this thesis.

Optical sensors also need not be confined to measuring appearance (and utilized for recovering geometry), it is only required that they detect of optical phenomena and that unfiltered data is available¹². LIDAR, RADAR, active illumination, stereo vision, thermal, and multispectral imaging are among the candidate sensors for this approach.

Several arguments can be made against this mostly qualitative analysis. First, the knowledge could be so obvious that it must either be hardly effective or already in common use. Second, it is possible that the variation in any one environment is so large that a single contiguous appearance space does not suffice. The former point is not likely in planetary environments as demonstrated by the lack of monocular vision-based modeling approaches. The latter is more enduring, environmental variation and

¹⁰ Technically, $\vec{n} = [f_x, f_y, -1]^T$ for most definitions of the gradient in computer vision. The terminology for appended vector and normalization are omitted to emphasize simplicity of relationship.

¹¹ The sun subtends an angular diameter of 0.53 degrees. It is a small area source, but it is possible to approximate it as a point source for simplicity.

¹² Availability of raw data is a non-trivial concern. Many commercial radar manufacturers, for example, limit the ability to query specific returns or low level information that would otherwise fit into this illumination framework.

uncertainty must be accepted in any application. Improvements from any technique can only be gauged probabilistically.

In light of these arguments, the thesis advocates experimental characterization of image-forming properties of domains to validate assumptions and gauge the variations in these distributions. Domain characterization also provides an ability to analyze the error of approaches new and old. Material frequency can be estimated from data in prevailing literature while unique reflectance functions can be found with calibrated imaging. Tessellated geometry in mines and lunar rock distributions are similarly suitable for characterization of effects like angular prominence of occlusion edges. The effects of sub-macroscopic geometry can be resolved with commodity range sensors at shorter, calibrated distances or using ultra-precise survey scanners. Natural illumination in the case of planetary space is often known (dark or repeatable) with high confidence such that characterization is redundant. Detailed methodology for characterizing the appearance of planetary spaces is presented in the following chapters.

Chapter 3:

Thesis Statement

This thesis addresses issues of data enhancement, optimality and presentation in planetary modeling by utilizing intensity images in the sensing process. The unique innovation of this research is the exploitation of constrained appearance with controlled illumination and estimation of material and geometric properties. This approach is made possible through understanding of optical domains and the parameters of image formation. Techniques developed herein enable generation of geometric models with quality and efficiency far beyond traditional ranging sensing methods.

This thesis asserts that planetary models are vastly improved with Lumenhancement – the targeted exploitation of surface appearance by incorporating intensity images, active illumination and material properties with range sensing.

Thesis Statement

This approach is coined *Lumenhancement* - from *lum* meaning “light” - because modeling sensors are overwhelmingly optical - and *enhancement* which describes the process of augmenting range sensing methods. Lumenhancement succinctly summarizes the core assertions of this work. First, that existing techniques can be improved, not replaced, with proper understanding of radiative transfer; and second that ideals of “gross generality” should not hinder adoption of targeted approaches that significantly enhance data quality when modeling is the principal objective of the mission.

Lumenenhancement in this thesis lies at the confluence of three topical areas: (1) **planetary modeling**, (2) **range and intensity image fusion** and (3) **active illumination**. Planetary missions, where current state-of-the-art modeling falls far short of necessity, serves as the catalyst for this research. However, the principles introduced in this work are broadly applicable to other domains. Active illumination is class of techniques for recovering information from intensity images championed by this thesis for relevance in these illumination-simple planetary environments. Lastly, fusion of intensity imagery from cameras with LIDAR ranging is the vehicle through which illumination-based techniques can address the issue of model enhancement. These themes are introduced below while relevant background is discussed in the following section.

Planetary Modeling. The primary application context is robotic planetary exploration, where domain relevance, economy and necessity compel advances in three-dimensional modeling. This domain represents a substantial segment of hazardous, yet essential, field environments where robotic operations have clear advantages over human ones. Moreover, models currently generated from human, surface and remote sensing techniques are of insufficient quality to meet the reconnaissance demands of these environments. This research characterizes the optical appearance of this domain by performing empirical analysis of material and geometric properties. Domain knowledge as well as survey of common applications grounds the development of new approaches and enables generalization of techniques developed for any member environment to others. Ultimately, performance in real planetary settings serves as the gauge of efficacy.

Range and Intensity Fusion. Direct sensing from robotic investigation is the only method of producing quality maps in remote planetary spaces; however, current modeling methodology produces poor returns on the cost of robotic deployment. The consequences are particularly egregious, including in terrestrial underground voids where inadequate methods are chosen over robotic mapping, often with “acceptable” risk to humans. Multi-sensor fusion can alleviate this problem by enhancing modeling beyond the sum of the individual sensor streams. In particular, range sensors, which measure 3D surface geometry and cameras which measure appearance have great synergy due to their complementary physics, common simultaneous usage and low cost. Mutual-information and cooperative data acquisition can enhance a spectrum of data modalities like range accuracy, acquisition speed, sampling density, reconstruction and scene understanding. However, the recovery of geometric information from intensity images is only probabilistically valid, and often fails in unconstrained, optically complex environments.

Active Illumination. Planetary environments have the unique attribute of simple natural illumination. This includes the total darkness of subterranean voids, polar craters and tidally locked bodies and environments that exist between darkness and direct (unscattered) sunlight such as skylights, asteroids, and lunar surfaces. The simplicity of natural illumination presents an opportunity to harness it for recovering geometric information in intensity images for fusion. Artificial lighting is required for imaging in dark environments, and thus may be explicitly constructed to recover scene information with calibration. Multi-flash photography and Structured Light sensing are just a few examples of techniques using this paradigm. In environments with simple natural illumination, like direct sunlight, the effects can be estimated in intensity imagery and reversed for the same purposes. In contrast, daylight surface environments hamper detection of artificial sources and preclude accurate estimation in images.

The use of “active illumination” in this thesis is distinct from prior work, and includes both the traditional meaning of geometric modeling with artificial lighting as well as the estimation of natural lighting so that it may be used, in effect, as a calibrated source. Varying modes of active illumination are explored in this thesis from single point-sources to wide-baseline flash photography and multi-spectral sources. Each of these lighting modes constrains image formation to isolate salient features such as depth, surface normals, occlusion boundaries, and albedo. These features are fused with range data for targeted enhancement of model properties.

3.1 A Model-Centric Approach

The objective of this work is to integrate intensity imagery and the concepts of appearance in the modeling process to enhance planetary map data. The duality of cameras and range sensors in data fusion is a promising solution to this problem. However, the effectiveness of camera-based techniques are inextricably tied to their use in permissible environments with valid assumptions. Fortunately, existing illumination and imaging infrastructure in the unique planetary domain enables easy adoption of these techniques to current robotic operations.

This synergistic relationship between fusion, appearance and domain is central to the idea of Lumenhancement. The profound improvements demonstrated by Lumenhancement are made possible only when these seemingly disparate topics are considered simultaneously. Surprisingly, this central consideration of models is a completely novel way of looking at robotic modeling.

Conversely, being a model-centric approach, this work does not delve into the traditional problems of robots in modeling. The only metrics of concern are improvements in model quality. Topics such as mechanism, SLAM, autonomy, communications and computing are left to prior work where they are covered with authority. Approaches presented herein are designed to be integrated with any number of robotic frameworks in a strapdown fashion and this fact is demonstrated in experimentation.

Chapter 4:

Background for Topical Areas

This section describes relevant background and prior art for each of the three topical areas, planetary modeling, and active illumination explored in this thesis. The thesis draws specific inspiration from these examples but novel contributions as they relate to appearance domains and models are contrasted in each of the categories.

4.1 Planetary Modeling

Given the remote nature of these spaces, scant prior work in planetary robots have resulted in missions to the intended application environments. Perhaps the only cosmic example with significant mapping capability¹³, the two Mars Exploration Rovers (MER) have operated on the red planet since 2003¹⁴. Each rover features a navigation stereo pair and panorama pair (PanCams) on a sensor mast and four smaller hazard pairs [Maki, et al. 2003]. Significantly, the rovers do not feature active range sensing such as LIDAR, which is distinct from many terrestrially-deployed planetary robotic testbeds such as *Scarab* [Wettergreen, et al. 2009], *Nomad* [Vandapel, et al 1999] and *K10* [Fong, et al. 2008]. *Spirit* and *Opportunity* have mapped over five kilometers of traverse each, of which 25% utilized automated terrain assessment from stereo [Maimone, et al. 2007]. Their mission has demonstrated the efficacy of vision-based modeling in the field, particularly multi-view techniques [Matthies, et al. 2007]. Geometric cues gleaned from images include point clouds and hazard maps from stereo vision [Goldberg, et al. 2002] and visual odometry for estimating robot egomotion [Cheng, et al. 2005].

The operation of MER produced several examples of tailoring domain knowledge to improve optical sensing similar to this thesis. The approach of [Willson, et al. 2005] quantifies and removes the optical

¹³ The earlier Mars Pathfinder rover, *Sojourner*, utilized a light stripe and stereo combination but could only calculate 20 points per image [Maimone, et al 2006].

¹⁴ As of 2012, one rover, *Opportunity* is still operational. *Spirit* became disabled in 2010.

effects produced by Martian dust particles on lenses. The MER PanCams, which are intended to identify mineralogic and photometric properties of surface materials utilizing spectral filters, are calibrated with analog materials and tested against a range of environmental BRDFs. Their stereo capability also provides point cloud geometric models, though these multiple vision functions are considered independent in the mission [Bell, et al. 2003]. Lastly, knowledge of material BRDFs in the form of the opposition effect was used to reduce false positives in the descent imagery of the MER lander [Cheng, et al 2006].

Notable terrestrial examples in planetary analog environments have also utilized characteristics of the domain for modeling. [Thompson, et al. 2008] examines the generalization of mapping as view sampling of features of interest. These features are not limited to geometry and can be geological or biological in nature, with application to finding life in the barren, planetary-like Atacama Desert. The segmentation of rocks using prior knowledge of texture, shape (raised boundaries) and shading from sun using ephemeris estimates for Martian imagery is explored in [Dunlop, et al. 2007]. [Vandapel, et al. 1999] documents sensing in the search for meteorites in Antarctica using the robot *Nomad*. Previously unfamiliar use of mapping sensors - stereo, LIDAR and RADAR – are characterized in planetary and white out conditions. Similarly, the work of [Pedersen, et al. 2008], characterizes performance of active range sensors in dark crater mapping. This thesis builds on prior planetary physics-based vision by performing characterization of the domain as well as the sensors. Techniques presented here also differ in their systemic nature and use for general geometric modeling as opposed to application specific problems.

Subterranean Modeling

In subterranean spaces, the work of Morris [Morris 2005; 2006; 2007] is particularly relevant to this thesis, as it investigates planning and autonomy for mapping in the underground domain described by Omohundro. Novel solutions are devised for error recovery in autonomous exploration, while maps are produced primarily as a byproduct of bootstrapping standard Simultaneous Localization and Mapping (SLAM) techniques. A method for integrating a mobile robot, retroreflector and theodolite for first autonomous survey is demonstrated on *CaveCrawler*. Resulting three dimensional models are presented as raw “point clouds” of range information or as a simple “mesh” of polygons generated from techniques such as Delaunay triangulation. Among the more effective methods of display presented are 2.5D birds-eye-view “relief” maps with contextual photographic, temperature and gas sensing information.

However, a systemic approach to filtering or data enhancement is not emphasized, and attempts to fuse, cross validate, or re-acquire with multi-sensor data are left for future work. Nonetheless, models created from field experimentation represent the state-of-the-art in underground modeling in terms of quality and comprehensiveness. The vast amount of experimentation and number of datasets generated from the research – some of which are utilized in this thesis work – are impressive. This thesis expands upon the modeling paradigm presented by Morris by innovating an approach where the express quality of the data is the focus of sensing

While this thesis could be considered the spiritual successor of the work of Omohundro and Morris, it distinguishes itself in several ways. This research is the first to quantify the subterranean (and the more general planetary) domain using the metrics of sensor physics. Unlike prior work, which emphasized domain knowledge as challenges to overcome, this work utilizes that knowledge for sensing enhancement. Moreover, domain taxonomy is employed to target and infer the performance of developments in environments which are beyond the scope of immediate experimentation.

Omohundro’s contributions (described in section 2.2) considered robot mechanical configuration for mapping underground voids using range sensing. Much of the design criteria concerned optimizing for sensor *coverage* from static viewpoints, a topic that is also explored in this thesis, but with mathematical rigor using sampling theory (see section 6.2). This work specifically rejects LIDAR-based range-sensing as a panacea for modeling and argues for a multi-sensor approach. The view of data enhancement and visualization advocated in this thesis also differentiates itself from that of Morris. Color information is not simply “painted” on range models, but is used to texture, interpolate and align data. Adaptive point displays are explored to combine the benefits of point clouds with that of meshing. Even realism and accuracy are not sacred; non-photorealism and texture replacement are exploited to enhance readability and awareness.

Localization and Navigation

Much of the prior mapping research described here have utilized SLAM for integrated autonomy and modeling, based on the seminal work of [Whyte, et al. 1996] and [Thrun, et al. 2000], among others. Traditional use of the words *mapping* and *modeling* in robotics context usually references these approaches. This thesis also utilizes variants of these approaches in parts of the experimentation framework. However, SLAM and navigation techniques are treated as black boxes and no significant novel contributions are made in this area from this work. Instead, this thesis **focuses on the sensing**

aspect of models and the utilization of all available appearance knowledge in modeling. The approaches here complement and assume the availability of state-of-the-art localization and navigation.

4.2 Range and Intensity Fusion

The fusion of range and intensity imaging sensors has been studied in depth for range filtering, interpolation, contextual understanding and visualization. Range sensing and intensity imaging are natural complements and arguments for fusing these sources to produce 3D models are compelling. Typical beam range sensors are single-detector and require mechanical actuation for scanning, resulting in low throughput and sparse readings¹⁵. Conversely, CCDs are solid-state, high throughput and high density but do not directly provide geometric information. Expensive LIDAR systems are already the de facto standard for 3D sensing in robotics; quality cameras are light-weight, low-power, inexpensive, and can greatly enhance the ability of LIDAR for little extra cost. This thesis advocates range and intensity fusion techniques for enhancement of planetary models. A brief survey of existing fusion research pertaining to the thesis follows.

Edge Localization. Perhaps the most studied concept in range/intensity fusion is the detection of boundary edges to filter range models. This is particularly beneficial in aerial mapping, where scanned LIDAR height data is rough and poorly localized due to sensor ego-motion, while instantaneous image capture retains crisp features. The work of Shenk and Csatho is an early success in deriving geometric cues from intensity imagery for application in filtering range models [Shenk, et al. 2002]. LIDAR digital elevation maps are fused with 3D terrain models recovered from multiple aerial images using human-directed stereopsis. The stereo data is dense, but often produces oversmoothing and matching errors, while the sparse LIDAR data must be interpolated¹⁶, but has a predictable height variance. Both sources exhibit edge inconsistencies. The innovation of the research is that intensity edge features are utilized from the images to clean edges in the fused heightmaps.

More recent related work [Holte, et al. 2008] have used the superior localization of intensity imaging to enhance extraction of gesture features in motion-blurred range video. Human gesture detection

¹⁵ Lensed range sensors like flash LIDAR are a nascent technology that promises high throughput through CCD-like silicon range detectors. Current models, however, offer lower throughput than that of most commercial point scanners with much lower range accuracy. The theoretical accuracy of these devices is limited by the use of flash sources which create cross talk and ambiguities.

¹⁶ The method actually calls for clustering of the LIDAR heightmap to produce elevation “patches,” which is equivalent to performing a nearest-neighbor interpolation.

requires both identifying differential edges in consecutive video frames and extracting geometric features from these edges. The latter can only be done using range space, while the former is accomplished rather poorly for several reasons. Oversmoothing of occlusion boundaries by linear interpolation during construction of the range image results in low gradients at range edges. High measurement variance exhibited by the flash LIDAR used in experimentation produces a noisy gradient in naturally smooth areas [Holte, et al. 2008], [Lindner, et al. 2008]. Lastly, fast motion of the subject creates motion blurs and ambiguities when captured by a slow exposure ranging device. By performing edge detection in intensity space, searching is quickly narrowed to regions of interest and resultant edges can be checked for false positives. Other contemporary edge-fusion research includes using Markov Random Fields, a probabilistic Bayes framework, for segmentation [Chang, et al. 2001].

Super Resolution. Upsampling is also a compelling argument for intensity imagery. Schenk's method re-parameterizes both data sources as sets of low dimensional surface patches. The application does not require minute 3D detail, and decimation has the benefit of macro-scale robustness and algorithmic and memory simplicity. However, high resolution imagery contains information about scene structure between range readings that cannot be deduced from pure interpolation or decimation of sparse LIDAR data. If image intensity correlates even weakly with range, then range information can be inferred from images beyond the information-neutral averaging process of interpolation. The process of correlating co-located range and intensity readings and inference of the relationship for interpolation is known as *super-resolution*. Super-resolution has potential for greatly increasing the coverage, density, and accuracy of 3D measurement with very low overhead.

A general model for fusing raw LIDAR and image data into super-resolution range images using a Markov Random Field (MRF) was explored in Diebel and Thrun's seminal paper [Diebel, Thrun 2005] (see also section 0). MRFs are undirected graphs that represent dependencies between random variables and have been used extensively in computer vision for noise removal, feature matching, segmentation and inpainting [Li 2001]. The popularity of the MRF stems from the ability to model complex processes using only a specification of local interactions, the regular grid nature of CCD images and the maximum likelihood solution requiring only direct convex optimization. In fact, it has been recently shown that MRF solutions can be computed on modern Graphics Processing Units (GPUs) for real-time speed ups [Vineet, et al. 2009].

Diebel and Thrun surmised that higher resolution intensity (color) data could be used to increase the range accuracy of interpolated points. In particular, it was noticed that low-gradient areas in the

intensity image correlated highly with true low-gradient (flat) surfaces beyond reconstructions from noisy, single-shot LIDAR readings alone. The results in a uniformly and sufficiently illuminated regular office environment are quite compelling. Cameras are able to turn LIDAR scans into dense range images with very low computational overhead. The ability of the method to smooth point clouds using areas of flat image information was convincingly shown, but the converse of enhancing a point cloud using image texture was not. The work of Diebel and Thrun has generated critical interest in range/image super-resolution, and notable extensions have proposed more expressive MRF models and feature detection, [Yang, et al. 2007], [Torres-Mendez, et al. 2007], [Gould, et al. 2008], [Harrison, et al. 2009].

The related work of [Mostafa, et al. 1999] proposed fusion of shape-from-shading reconstructions with LIDAR in a neural network framework. Shape-from-shading methods recover the per-pixel surface normal in intensity images by estimation of illumination and reflectance functions in the geometric appearance model. If the estimates of illumination and reflectance are accurate, SFS methods are much more effective at revealing true geometry than strictly probabilistic methods (such as edge finding). In Mostafa's approach, the image depth reconstruction is fed into a neural network which encodes the expected error of the image estimates with respect to the range data (assumed to be perfect). This expected error is then used to correct points on the SFS model where there is no corresponding range value. The resulting corrected SFS model is naturally dense.

While this Mostafa generates physics-based geometric image features and predates the unconstrained image fusion of Diebel, it is not without several fundamental limitations. Firstly, iterative sample-neighbor interactions cannot be encoded in a feed-forward neural network, meaning that the resultant surface may neither be smooth nor consistent. Moreover, the model does not account for sensor or interpolation uncertainty. Training of the neural network is slow and prone to overfitting. Most importantly, there is an assumption that it is possible to train a function that will (correctly) determine the error of SFS points without corresponding range values using locations in the scan with these values. Such an error function would be highly dependent on sampling specifics, computationally intractable for complex scenes and undeterminable in the general case.

Feature detection. Intensity and color information, like geometry, is an independent attribute of objects that can be used for identification. The high density and multi-channel nature of color imagery once again shines in feature detection, as it is often more feature rich than its sparse 3D counterpart. Using vectors of both range and intensity features often improves discriminativity and detection rate, while

matching of geometric features may be sped up with preliminary search pruning in intensity space and vice versa.

The work of Lee and Stockman demonstrates assembly line detection of partially occluded objects [Lee, et al. 1998] using fused range and intensity imagery. Their method generates features called wing primitives which are view-dependent 2.5D extensions of polyhedral objects and line drawings [Baumgart 1975], [Malick 1987] to range images. Polyhedral line drawings are a framework for reasoning about of 2D projections of 3D objects by simplifying shape as collections of faces, edges and vertices. The types of vertices in the polyhedral world are deterministic and representation of objects reduces to edge labeling and vertex selection from a dictionary. Objects are detected by matching the order and types of vertices with a known template. However, as projection is a lossy transformation, inferring 3D structure from edge order produces ambiguities which are unacceptable with a large catalog of objects. Wing primitives employ range information to resolve these ambiguities, while more discriminative intensity information is used for segmentation and edge extraction. Use of both range and color data also increases the probability of producing matches even though part of the object may be occluded. Similar work utilizing augmented intensity and range features include that of [Baltzakis 2003], for landmark detection in occupancy grid based SLAM, [Stipes, et al. 2008] for neighbor-matching in Iterative Closest Point (ICP) alignment of overlapping range models, and [Gould, et al. 2008] for indoor object detection.

The use of range and intensity modalities for online learning has also been explored with great success. Stanford's winning entry in the 2005 DARPA urban challenge, the robot Stanley, uses range and color fusion for high-speed terrain classification. Short LIDAR sensing horizon at high speeds greatly reduced the robot's ability to avoid obstacles; range data alone was insufficient in detecting obstacles far enough away to maneuver around them. Cameras with targeted fields-of-view augmented the obstacle avoidance ability of Stanley by conservatively overclassifying obstacles at range [Thrun, et al. 2006]. Planned trajectories avoid these obstacles with margin while fine maneuvering is made with short range LIDAR. The accurate 3D models produced up close are then used to automatically tune image parameters. This enables adaptation to new drivable non-paved terrain such as grass and dirt and provides robustness against illumination changes in imagery. More recent application to automated learning is presented in work of [Mastin, et al. 2009], which uses optimization of the mutual information between range and intensity to learn global alignment of unregistered aerial imagery and elevation maps.

Visualization. Recent cost reductions and accuracy increases in both cameras and ranging sensors have led to marked advancements in the field of computer graphics. Rendering, the presentation of scene models under differing viewpoints, requires precise knowledge of geometry, material and illumination, necessitating some coordination of range and color acquisition [Waggershauser 2005]. This information is often estimated for rendering of real scenes by texturing single-view photographic color onto raw geometry [Fruh, et al. 2003], [Morris, et al. 2007]. Texture mapping is fast and simple, but fails to produce *believable* results when viewed off-perspective, except under ambient illumination. Techniques beyond texture mapping are generally divided into photorealistic and non-photorealistic methods.

Photorealistic rendering has demonstrated high potential in digitizing immersive, multi-perspective scenes for CGI films, games, and 3D displays. Moreover, the advent of mixed range and color sensors has made large scale digitization a possibility in the last decade. However, while producing stunning, archival quality results, collection of photorealistic scene data requires prohibitively expensive and complex gantries [Levoy, et al. 2000]. Consequently, much research has been directed towards image-based rendering, which infers the external appearance of the scene from a four dimensional light-field [Pulli, et al. 1997]. This light-field can be captured by taking calibrated images about a hemisphere of views using setups like camera arrays [Wilburn, et al. 2002], plenoptic cameras [Ng, et al. 2005], and coded aperture masks [Veeraraghavan, et al. 2007]. However, drawbacks of these image-only methods include narrow angles of sampling, poor sampling density and poor geometric reconstruction. Debevec proposes a hybrid technique of collecting low resolution range models coupled with an environmental radiance map captured using a mirrored sphere contained in the scene [Debevec 1998]. The results present sufficient realism in rendering synthetic objects placed in a pre-captured scene, but cannot extrapolate a full set of realistic views. More recently, hybrid techniques have used co-registered camera and end-user LIDAR systems with GPU acceleration to accurately build geometric models in real-time while enabling freedom of movement during full light-field capture [Todt, et al. 2005], [Waggershauser 2005].

Contrary to imaged-based rendering, non-photorealistic rendering (NPR) techniques do not attempt to reproduce the appearance of a scene under novel views. Instead, NPR techniques convey the most information to the viewer by exaggerating color, texture, occlusion edges or illumination [Gooch, Gooch et al. 1998; 2001]. In particular, NPR has found wide-spread acceptance in the gaming industry, with examples including cartoon [*Borderlands*, Gearbox Software 2009] and 1950s art-deco rendering [*Bioshock*, Irrational Games 2008]. Mostly recently, interest in non-photorealism has renewed as a

potential method for improving situational awareness in human control of equipment with applications to mapping, combat and mining [Summers, et al. 2005], [Winnemoller 2006].

In the aforementioned fusion work, the implicit assumption is that intensity or color discontinuities correlate with range discontinuities. This is tenuous at best, as evidenced by the vast prior art in image-based shadow removal, illumination invariance, and material invariance [Finlayson, et al. 2001], [Maxwell, et al. 2008]. Yet, this assumption features in the vast majority of range/intensity fusion research. Diebel's method, for example, biases fronto-parallel planes, which are commonplace in the office application environment presented, but rare in field robotics. Even in cartoon rendering, edges in a single-view image may correspond to albedo, self-shadowing or perpendicular normals, and not discontinuity in range. Generating 3D geometry from a 2D projection is an ill-posed problem. In general, material properties cannot be separated from appearance without *a priori* knowledge [Shenk, et al. 2002].

This thesis extends these range and intensity fusion techniques to a subset of outdoor environments by utilizing domain knowledge. Intensity imagery is a necessity for exploratory robots, particularly underground, but it is often not used in 3D modeling. This work asks: **if imaging is a necessity and most imaging requires active illumination, why not utilize *intelligent* illumination?** This work specifically solves the issue of under-constrained image reconstruction by utilizing calibrated light fields and estimating material reflectance and geometric properties to constrain image formation. Changes in image intensity can be *constructed* to correlate with range or material changes, to a high degree of probability.

4.3 Active Illumination

Actively illuminated sensors use calibrated, artificial light to measure scenes. Natural illumination, whether complex or nonexistent, often precludes visual, geometric scene understanding. Use of artificial lighting clarifies appearance by controlling one or more factors of under-constrained image formation. Actively illuminated sensors include LIDAR, RADAR, and structured light – which directly measure geometry – and intensity imaging – which measures a product of material and surface properties.

This research utilizes a combination of Time-of-Flight light sensing and active intensity imaging for planetary modeling. While light ranging is well established, active intensity imaging and the fusion of

both for outdoor modeling is a nascent field first addressed by this thesis and preliminary work of the author [Wong, et al. 2009], [Wong, et al. 2012].

A rich body of prior work (discussed below) exists in the use of actively illuminated intensity imaging for indoor purposes. Undoubtedly, the largest obstacle to using active illumination in general outdoor scenarios is the existence of brilliant illumination from the sun. Sunlight irreversibly suppresses the signal to noise ratio of the artificial source and even dazzles many near-infrared LIDAR sensors. Subsequently, much research has focused on outdoor illumination estimation [Finlayson, et al. 2001], [Maxwell, et al. 2008], effectively utilizing the sun as a “calibrated” light source. However, even with perfect ephemeris and registration with the sensor, secondary, but significant contributions from the sky and interreflections from the scene cannot be unambiguously separated from direct sunlight.

Many planetary environments are prime candidates for use of active illumination for the obvious reason of natural darkness. However, total darkness (while advantageous) is not an absolute requirement for using these techniques. Bounded planetary environments with point sources and negligible scattering atmosphere, such as sun-lit craters, or ephemeral sources, such as head lamps in mines, are amenable to active illumination. For example, the contribution of the natural light can be removed by exploiting superposition property of light and subtracting a background image. This section introduces some of the prior research in active intensity imaging with potential for use in planetary modeling. The corpus of active illumination techniques can be divided according to the complexity of the light source modeled. Figure 8 illustrates some common light source setups, of which 1-3 are utilized in this thesis.

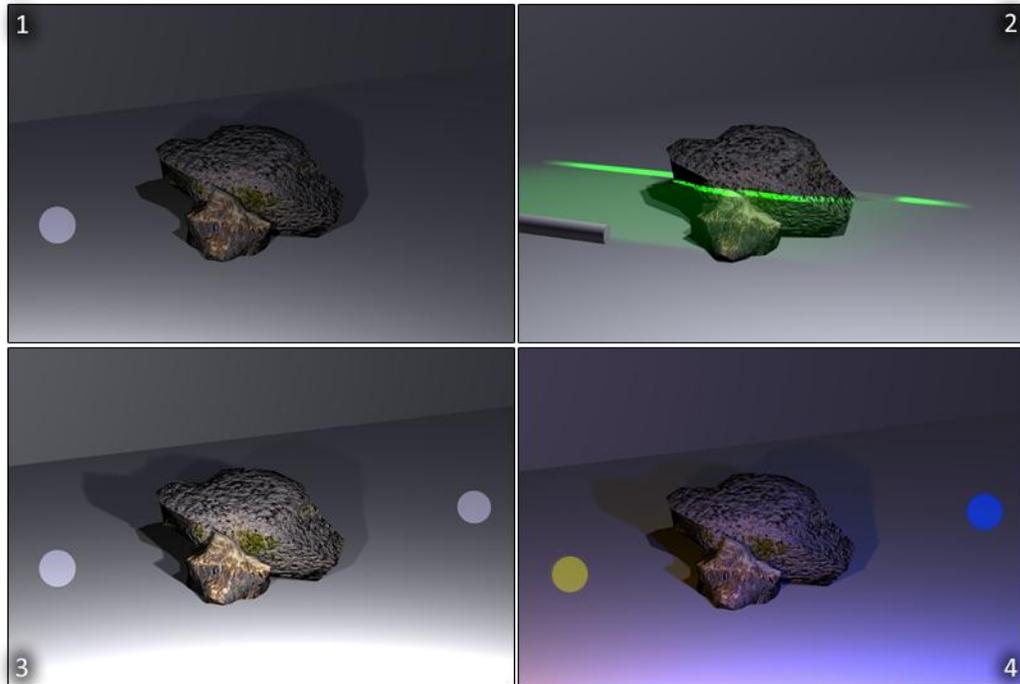


Figure 8. Sample Artificial Light Source Configurations. A simulated rocky scene is illuminated by (1) a single near point source, (2) a directional linear source, (3) multiple near point sources, (4) multiple, spectrally-distinct sources.

Shape from Shading. The shading of an object correlates with geometry in the form of surface normal direction. Humans, in particular, are attuned to perceiving geometry from shading information, a fact exploited by artist and painters [Horn 1970], [Zhang, et al. 1999]. However, the equations of image formation demonstrate that appearance is an often inseparable product of illumination, geometry, and material. Nonetheless, reconstruction of 3D shape from 2D imagery is a classical problem in computer vision. Automated generation of accurate geometry from a single image has immense ramifications for modeling and digitizing the world from pictures. Unfortunately, this is a profoundly difficult problem for reasons mentioned previously. Early research into Shape from Shading (SFS) often restricted multiple parameters of appearance, in the form of assumptions, to make the problem tractable. The seminal thesis work of Horn surmised the relationship between geometric gradient and surface reflectance (manifested in the image) as a non-linear first order partial differential equation in two unknowns [Horn 1970]. The solution to shape recovery was numerical integration of five equivalent ordinary differential equations along characteristic curves from an initial known location. Horn's work assumed known uniform (diffuse) surface reflectance, source type, source direction(s), smooth surfaces, and initial curves. However, even with these assumptions, there is still ambiguity in whether a surface is convex or concave. Thus, experimentation was limited to simulated convex objects and an application to facial

nose detection where absolute structure obtained was less important than differences between subjects.

Successive research has extended the Shape from Shading of Horn with ability to handle subsets of more complex or unknown lighting, non-convexity, specularity, non-uniformity, inter-reflections [Nayar, et al. 1990; Yang, et al. 1997], and cast shadows [Stork 2006; Smith, et al. 2006] among others. The gamut of work can be further classified according to the type of solution-finding technique used. These techniques include energy minimization, surface propagation, local and linear methods [Zhang, et al. 1999], of which minimization and local methods are of particular interest in this text. Minimization methods treat the problem as an optimization of recovered shape with constraints such as local smoothness, unit brightness, and integrability. These methods have arguably emerged as the most robust in generating physically-valid shapes with minimum depth error as compared to ground truth. However, global minimization produces significant oversmoothing of surface detail and initialization of boundary conditions is a difficult problem in itself. At the opposite end of the spectrum are local methods which approximate neighborhoods at the pixel level with known geometry such as a sphere. Slant and tilt normal angles are estimated by inverting the image formation equations at each pixel using the geometric assumption. Local methods produce globally inconsistent models with poorer accuracy, but they preserve local surface features well and do not require boundary detection [Lee, et al. 1985], [Ferrie, et al. 1989].

The common theme in SFS research is that illumination is known or can be estimated from image cues. Furthermore, illumination is implicitly required to be expressible as a linear super-position of simple sources. While these illumination requirements are not easily met in general environments, SFS algorithms have high relevance to planetary modeling. Ideal lighting conditions can be expressly constructed in dark environments for the purpose of shape recovery. Moreover, the prevalence of Lambertian-like materials enables the use of the simplest and most robust algorithms.

Separable BRDFs and Color. The spectral composition of illumination is a fundamental component of image formation. When contributions to the reflectance functions of materials are linearly separable, knowledge of light color can be leveraged to recover details of scene appearance. Schafer's dichromatic model of reflection asserts that the total reflectance function of a surface can be decomposed as additive diffuse and specular components [Shafer 1985]. Most materials further obey the Neutral Interface Reflectance (NIR) property, which models the specular component as a perfect spectral mirror

of the light source. The appearance of object color is explained by spectral shifts caused solely by diffuse interaction between the incoming light and material pigments [Lee, et al. 1990]¹⁷.

The work of [Zickler et al. 2006; 2008] uses the NIR assumption to remove specularities in images lit by a single point-source of known color. The key idea is rotation of the image's RGB pixel space so that one of the basis vectors (R, G, or B) aligns with the source color. This rotation transforms tri-channel RGB space into a colorspace of two diffuse and one specular channels, called the SUV space. The specular channel (S) is the component parallel to the source color and therefore encodes all of the specular reflection along with some unknown portion of the diffuse. Discarding the S channel produces a 2 channel image consisting entirely of diffuse reflectance. Unlike heuristic, illumination-invariant colorspace such as LAB or r-g chromaticity, which are projective, coordinate rotation is a linear operation and preserves shading information. Thus, the power of the SUV transformation is that it enables vision algorithms requiring Lambertian reflectance to run natively on non-Lambertian data. In particular, Zickler demonstrates that the SUV transformation can increase the accuracy of surface normal recovery on highly specular objects. Conversely, specular areas or objects can be identified by analysis of the S channel. A degenerate condition occurs when the illuminant and scene have identical colors. In this case, the entirety of the image's spectral power is distributed in the S channel and no diffuse information will remain after removal. While this scenario is extremely rare in practice, the signal to noise ratio (SNR) of the diffuse components is adversely affected by the relative spectral power distributions of the scene and the source.

Separable BRDFs are also featured in [Narasimhan, et al. 2003], which derives a class of photometrically invariant features from multiple images of the scene. Unlike the SUV transformation, the work makes no assumption of separable diffuse and specular reflectances, but requires that the BRDF can be written as a sum-of-products form of purely geometric and material functions. Amazingly, many analytic BRDFs are factorable in this way. By modulating either geometry (light position, correlated pixel locations in a moving scene) across several images while fixing material properties (source color, RGB filters) and vice versa, a feature transformation can be constructed for each pixel that is invariant to either material or geometry changes. Narasimhan proposes using the determinant of block sub-matrices in an augmented observation matrix of a pixel. The matrix consists of a single pixel's value across changing material images (R, G, and B spectra) in the rows and changing geometry images (light positions) in the columns. The choice of sub-matrix governs whether all the geometric or material terms are implicitly canceled in

¹⁷ Notable exceptions to the NIR model are shiny metals.

the determinant arithmetic. The resulting feature value is dependent only on the property that was not canceled and is densely generated for every pixel in the scene. In particular, the *geometry* invariant form is significant for its use in detecting and discriminating unknown materials without requiring spectroscopy. It requires at minimum two tri-channel RGB images of two sources illuminating the scene. The 3x3 observation matrix used in the paper handles materials with 2-term BRDFs; however, the technique can be easily generalized to more terms.

Multi-Source Methods. As a result of the close coupling of lighting direction and surface normal in reflection (a dot product relationship), multiple calibrated light sources (or a single motional source) can be used to understand geometry in a variety of ways. One such example is Narasimhan's aforementioned photometric invariants, which were partially generated by acquiring multiple images of the scene under differing illumination positions.

A particular class of methods assumes known (or estimable) source positions in three or more images to recover three-dimensional surface normals. These methods, collectively called photometric stereo, are generalizations of shape-from-shading (SFS) and were first explored for the Lambertian case by [Ikeuchi, et al. 1979] and [Woodham 1980]. The problem can be recast as solving a matrix equation of three unknowns, which are the three components of the normal vector at a pixel. The solution can be obtained by inversion of the 3xn matrix of known source positions and pre-multiplying with the pixel brightness value. If the number of source positions is three or greater, the over-determined system estimates normals unambiguously, unlike in the SFS case. While in theory, these methods will recover the normal at every pixel exactly, they are quite vulnerable to measurement noise in practice. The resultant normal vector field, while perceptually correct, often cannot be integrated directly if it is not conservative to numerical precision. This noise grows as the angular separations of the sources in the scene are reduced. Like SFS, these methods also suffer from inaccuracies due to shadows and interreflections. However, in contrast, the use of a calibrated mount of three or more sources is operationally infeasible for many mobile applications. Recent extensions of photometric stereo methods have made them more robust to unknown source locations [Hayakawa 1994], complex reflection and shadows.

Multiple sources, in the form of multi-flash photography, can also be used to specifically detect occlusions and object boundaries. These "depth edges" are distinct from changes in material and lighting, but are indistinguishable from both in general imagery. Raskar's multi-flash method exploits the appearance of cast shadows to segment foreground occluders from the background [Raskar, et al.

2004]. The main idea is that angular separation of a light source and the camera enables the imaging of shadows, but the same shadows will not be visible under different illumination parameters. A single image is taken of each source solely illuminating the scene. As long as every pixel is visible in at least 1 image, an unshadowed reference image can be generated by storing the max pixel values across all images. All source images are then represented as per-pixel fractions of the max image. Searching for large negative steps in the ratio images along epipolar lines between the source and camera center will detect the depth edges. At least two sources are required for single-dimension boundary detection, and three sources are required for full two-dimensional detection in the image plane. The angular separation of the sources (baseline) determines the valid imaging volume. Depth edges beyond this range do not produce visible shadows and edges too near produce detached shadows. The method is surprisingly robust for its simplicity; only large specularities produce false positives, but can be partially remedied by conducting a sanity check amongst multiple images or SUV specularly removal.

Multi-flash imagery can be used in conjunction with albedo estimation or geometry invariant features to validate material edges. Moreover, the high spatial resolution of CCDs enables pixel boundaries detected in this manner to clean and localize interpolated depth maps from LIDAR.

Structured Light is a method for direct range-finding in intensity images using image-based triangulation of pixels in a light pattern projected onto the scene. Often this light is a high-intensity coherent laser “stripe” (see Figure 8 – (2)) that overwhelms the reflectance characteristics of the surface in a narrow band to ensure detection. While this is not a true intensity imaging technique, the use of “active illumination” and the resultant geometric measurement make this a promising technique for planetary mapping.

The work treats structured light as a direct range sensor to which Lumenhancement with intensity images can apply. There are several advantages to using structured light as the primary sensor compared to time of flight methods. Programmable light sources require no mechanical actuation, unlike beam based LIDAR. Moreover, there is no constraint on the sample order or fixed resolution of parallel methods like flash LIDAR, a fact exploited in this thesis for the development of a new sensor. However, in the view of this thesis, the similar resolution and physics of structured light to monocular vision make this technology slightly comparably suitable for fusion than LIDAR range sensing.

The active illumination methods discussed in this section are well-adapted for planetary use with little modification. While the primary innovation of this thesis is the utilization of active illumination methods

in intensity image to improve LIDAR range modeling, the mutual synergy of both sensing modalities should not be ignored. The availability of sparse LIDAR data, for example, greatly reduces the complexity and uncertainty in SFS. Concavity, boundary and source direction can be ascertained with accuracy beyond current image-based estimation techniques even with poor LIDAR data. Local SFS techniques, which produce asymptotic distortions in noisy data, can be regularized with *a priori* knowledge of depth such that they are both volumetrically consistent (like global techniques) and still feature preserving. Even attached specularities are easily removed in multi-flash imagery by validating with the depth gradient. The use of preliminary range data to improve intensity image techniques in turn improves the final range estimate. This is *the* compelling motivation for sensor fusion in underground application: the end result is unachievable in the sum of its parts and unachievable in other outdoor domains.

4.4 Range Sensing Technologies

This research primarily addresses time of flight or frequency modulation LIDAR ranging techniques. However, triangulation-based ranging has been used to great effect for surface and indoor mapping. Triangulation utilizes the disparity of points in multiple perspectives of a scene to generate depth, which is inversely proportional to disparity. The offset transformation between these perspectives is called the baseline. Increasing the baseline enhances range accuracy, but also results in smaller overlap for geometric measurement.

Stereo Vision and Structured Light are popular implementations of triangulation. Stereo vision uses two cameras to image the scene and statistical matching algorithms to produce the disparity map. Stereo requires only simultaneous capture of two images and results in the fastest geometry generation. Acquisition often utilizes natural illumination and thus also consumes the least power. However, the correspondence problem is computationally expensive across images and prone to scene-dependent error. Structured light uses active illumination to paint the scene unambiguously and a monocular imager to detect the projection of the illuminated point. There are several methods for resolving the position of scene points [Salvi, et al. 2004] with “codes.” These codes require the capture of between $\log n$ and n images, where n is the number of sample points (pixels), for unambiguous recovery. Despite increased data, the approach is computationally simpler than stereo vision. Other, graph-based approaches are capable inferring the disparity map from a single image, but have many of the same problems as stereo correspondence.

While many of the improvements in this thesis apply readily to triangulation-based range measurement techniques like stereo vision or structured light, the characteristics of LIDAR are particularly advantageous in underground modeling and fusion with intensity imagery. Some of these advantages are discussed below to highlight the inability of triangulation-based techniques alone to match fused LIDAR and intensity imagery in addressing the issues of underground modeling.

Dark Operation. LIDAR is well suited for dark environments. Infrared lasers are scattered less by atmospheric particles and produce higher signal to noise at range than visible-light CCDs. This feature of LIDAR makes it advantageous for fusion with intensity imagery over triangulation techniques, which generate measurement using the same physical processes as imagery and thus exhibit the same noise characteristics. Stereo in particular has problems integrating with active illumination. Correspondence techniques rely on uniform appearance of points across multiple views, an assumption which holds only strictly for highly-textured, convex Lambertian scenes under ambient illumination. The use of active illumination, such as isotropic point, multi-flash, and multi-colored sources often creates harsh shadows, changing surface shading and spectral shifts in appearance. While these cues are often good local approximations of shape or material, global consistency, such as that offered by LIDAR is required to bound estimation errors. Stereo reconstruction, which has demonstrated critical success for low-power, mass-sensitive planetary surface mapping, is susceptible to these issues and performs less desirably in dark modeling [Pedersen, et al. 2008].

Range robustness. In theory, LIDAR range resolution is invariant to range value (Equation 3.1). In practice, LIDAR measurement accuracy varies weakly with range, with the major sources of error being photon shot noise and charge conversion noise in the detector. Both these errors increase slightly with dwell-time [Hussmann, et al. 2008], though they are easily mitigated with signal processing. Triangulation methods, however, which rely on a finite baseline between views with constant lateral pixel density in the imagers. The quantization of projected area (pixelization) results in range resolution decreasing and error increasing proportionally to the square of range (equation 3.2) [Rankin, et al. 2005]. Though, recent developments in stereo algorithms have honed robust sub-pixel estimation to reduce this effect [Stein, et al. 2006]. By same principle, triangulation produces ample lateral density in the near field - where density is arguably less important - and sparse measurement at range. While actuated LIDAR scanners also exhibit a finite angular resolution for sampling at range and non-zero beam divergence, near-field sampling can often be reduced for significant speedup.

$$\Delta r = \frac{c \cdot \tau}{2} \quad (4.2)$$

LIDAR Range Resolution. The minimum discernible change in range (Δr) is equal to the speed of light (c) times half the pulse width (τ).

$$\Delta r = \frac{r^2}{b \cdot f_{lens}} \Delta d \quad (4.3)$$

Triangulation Range Resolution. The minimum discernible change in range (Δr) is equal to the square of the range (r) divided by the baseline (b) and focal length (f) times the left-right disparity (Δd).

Erroneous measurement. LIDAR ranging measures the period between transmission and detection of light signals reflected off surfaces. The range to a surface is a straight forward linear proportion to the Time-of-Flight (TOF) of the light signal and the sensor is well-approximated by a pinhole. Unlike triangulation sensors, the quality of TOF measurement does not depend on the appearance of surfaces, only the that the strength of the return is sufficient for detection. Stereo vision relates disparity of identical pixels in multiple views - the correspondence problem - to depth. Incorrect matches often result from differing regions exhibiting similar textures, while unmatchable pixels occur in low texture, specular, or cast shadowed regions. Both these problems result in artifacts in the final depth estimate and incorrect calibration can magnify these errors. The practical resolution of a stereo model may be much lower than a LIDAR model for particularly difficult scenes. As the appearance of the scene strongly influences the quality of reconstruction, stereo vision is less appropriate for error-bounded, survey-like modeling. Structured light produces cleaner models than stereo vision in unknown environments. The use of single-point measurements or coded light patterns generates clear matches in the image. However, structured light sensors are still susceptible to scene appearance. Subsurface scattering causes poor localization of the measurement point and highly specular surfaces may produce ambiguous matches through multiple reflection (known as mixed pixels) [Levoy, et al. 2000]. Like stereo vision, the disparity between transmitter and detector can cause shadows in occluded regions.

LIDAR measures only range and in some cases single-frequency reflectance. In contrast, stereo vision and structured light can both generate models with full spectral color (under certain illuminants). With LIDAR systems, color information can only come from co-registered cameras which must be actively illuminated underground. Thus, it can be argued that if actively illuminated monocular imagery is necessary in addition to LIDAR, structured light triangulation should be used instead to generate a second, high density, absolute geometric observation. Indeed, the largest differences between point-

source illuminated imagery and structured light is the complexity of the programmable light source and the type of geometric information generated from the imagery. Most of the techniques described here can be implemented on a co-registered LIDAR and structured light system with great effectiveness.

While there are unique merits to any combination of sensors working together, there is perhaps a point where opportunity cost exceeds the gain. This thesis advocates the use of monocular imagery with one or two point sources as a middle ground between system simplicity and accuracy. Most underground robots require illuminated imagery for documentation or operator oversight; however, none of them require complex programmable light sources like those needed for structured light. Thus, intensity imagery can be bootstrapped to existing infrastructure, such as underground robots or even human-driven mine vehicles, with minimal invasiveness. With simplicity and adaptability, intensity fusion techniques have great potential to be game-changers underground.

Other arguments against fusing triangulation-based sensors with LIDAR include the introduction of erroneous measurements to pristine LIDAR models in difficult environments; complexity in calibrating for zoom imagery; and sensitivity to decalibration which can occur with frequent vibration in rough terrain.

Chapter 5:

Characterization of Planetary Appearance

This section quantitatively describes planetary appearance through exhaustive experimentation. Two parameters of image formation - material and geometric distributions – are characterized with physics-based measurement in representative environments. The third parameter, illumination, is arguably trivial to characterize, generally known, and is viewed as an exploitable (and changeable) entity. Furthermore, light-complicating media, such as dust is assumed to be negligible in this work. Thus, empirical characterization of light and atmospherics in the domain is not considered here and instead left to existing and future literature.

5.1 Analysis of Material Properties

The appearance of a scene under illumination is governed largely by the Bidirectional Reflectance Distribution Functions (BRDFs) of materials found in the scene. BRDFs describe the transmission of light as the reflection of incoming light rays to outgoing rays. This occurs in the hemisphere normal to every point on surface (Figure 9). As rays can be parameterized in two spherical coordinates, the BRDF is a 4-dimensional function. However, this is an idealization, as light does not truly “bounce off” immediately and singularly at each point. Subsurface scattering reflectance functions (BSSRDFs) more generally describe radiometry both at the surface and interior of materials in the formation of appearance. Yet, even these do not factor temporal and spectral effects like phosphorescence and fluorescence. Ultimately, mathematical intractability precludes consideration of many complex but fortunately rare phenomena. In this sense, the BRDF is the principal unique property of materials in computer vision.

The BRDF, along with albedo and geometry are sufficient to describe most non-translucent and non-radiant scenes - and the vast majority of planetary environments. Even many types of subsurface phenomena in the domain of the BSSRDF can be approximated to high fidelity as a BRDF and first order Taylor expansions [Jensen, et al. 2001]. Knowledge of environmental BRDFs is crucial to understanding and exploitation of appearance, as well as probabilistic reasoning about the effectiveness of such computer vision algorithms.

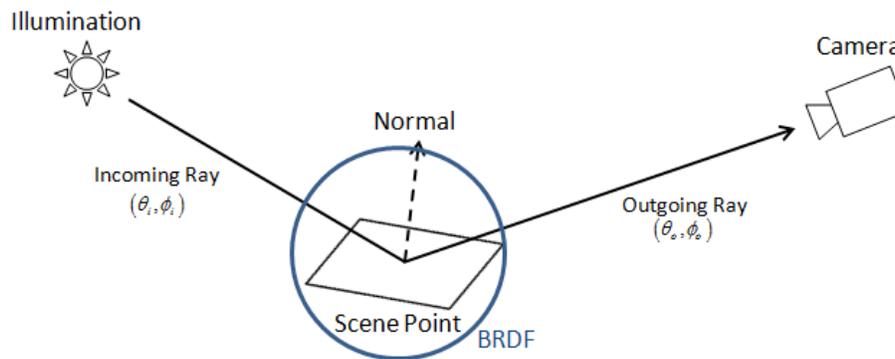


Figure 9. Role of the BRDF in the Geometric Appearance Model. The BRDF is a material-specific mapping of reflected light based on the direction of incoming and outgoing rays.

While the characterization, measurement and utilization of BRDFs has mostly been the realm of computer graphics researchers, the knowledge and approach to material analysis has profound significance for field mapping and imaging. In a classic example, stereo vision algorithms rely on establishing correspondences in images taken from differing viewpoints using similarity criteria. Such criteria are confounded by specularities, which are drastic changes in scene intensity dependent on view point. Thus, intensity features are only strictly valid for Lambertian surfaces unless knowledge of material properties and scene illumination are available [Wohler, et al. 2008]. Non-Lambertian deviations are an ongoing problem in planetary mapping and discussed in the work of [Nefian, et al. 2009], who have used stereo in the form of long baseline Lunar orbital imagery. While the moon appears matte and uniform, it is covered by a layer of regolith - a type of pulverized rocky dust - which acts as an amalgam of tiny mirrors exhibiting strong backscattering [Hapke 1993; 1998]. Shape-from-Shading (SFS) is another class of vision algorithms strongly affected by material reflectance. SFS, which infers depth and surface normal from intensity, is explored in this thesis for reconstruction. However,

classic SFS algorithms must assume both known BRDF and albedo values *a priori* for accurate reconstruction.

It is suspected that diffuse materials dominate many planetary and underground environments. Few people would describe the lunar terrain or a lava tube as “shiny.” This fact makes these domains of prime interest for utilizing physics-based imaging techniques that may not be applicable for field robots in general environments. However, few materials are perfectly Lambertian and most underground materials are an amalgam of rocks, minerals and dust, all with unique reflectances – some, like silica, are known to be quite specular in the micro-scale. Some diffuse objects, like the moon, are markedly non-Lambertian [Oren, et al. 94; 95]. It is principally useful to survey the types of materials that may be commonly encountered in a domain and to determine the degree to which these materials can be represented as Lambertian or other simple, invertible BRDFs. Accurate representation with invertible, analytical BRDFs is the primary driver of vision algorithm performance, regardless of whether a material is qualitatively diffuse.

While there are many existing databases of experimentally-derived BRDFs, covering a gamut of materials [MERL: Matusik, et al. 2003] [CURET: Dana, et al. 1999], there has been curious little interest in the vision community for examining planetary materials in this manner. This work extends the body of known BRDFs through experimental measurement of commonly encountered planetary materials. These results are used to quantify the optical properties of domain statistically, delineate different constituent environments within the domain, and to reason about the targeting of vision-based enhancements. Furthermore, material appearances captured in this experiment will also be used in a number of novel detection and rendering techniques discussed later.

5.1.1 Introduction to BRDFs

The BRDF, $f_r(\omega_i, \omega_o)$, is simply the ratio of light energy for all possible combinations of incidence and emergence, as seen in equation (5.1),

$$f_r(\omega_i, \omega_o) = \frac{dL_r(\omega_o)}{dE(\omega_i)} = \frac{dL_r(\omega_o)}{L(\omega_i) \cos \theta_i d\omega_i} \quad (5.1)$$

where $\omega_i = (\theta_i, \phi_i)$, $\omega_o = (\theta_o, \phi_o)$ are incoming and outgoing rays relative to the surface normal and their spherical coordinate parameterizations, L is the radiance and E is the radiance. Thus, the BRDF is a continuous, differential quantity.

Exact analytic forms for the BRDF of most materials do not exist, however several analytic BRDFs are grounded in physical models of radiant transfer. These tend to approximate common materials well but are mathematically complex. Physical BRDFs are required to obey several constraints: Positivity, Helmholtz reciprocity and Conservation of Energy. As a ratio of radiances, the BRDF must always be positive or zero valued.

$$f_r(\omega_i, \omega_o) \geq 0 \quad \forall \omega_i, \omega_o \quad (5.2)$$

Helmholtz reciprocity describes the symmetric nature of light by equating the scene radiances when transposing incoming and outgoing (view) light distributions. A gross simplification in layman's terms is "if you can see me, I can see you."

$$f_r(\omega_i, \omega_o) = f_r(\omega_o, \omega_i) \quad \forall \omega_i, \omega_o \quad (5.3)$$

The conservation of energy simply states that the reflected radiation (not including self-radiance) cannot be greater than the irradiance. Though objects may be self-radiant, the BRDF encodes only the interaction of reflected light. This is integrated over the hemispheres of incidence and emergence.

$$\int_{\Omega} f_r(\omega_i, \omega_o) \cos \theta_o d\omega_o \leq 1 \quad \forall \omega_i \quad (5.4)$$

An additional property of surface isotropy is assumed (and required) in this work, as the sensor can only detect 3 degrees of freedom. This states that if the material is rotated azimuthally, there is no perceptive change.

$$\begin{aligned} f_r(\theta_i, \phi_i, \theta_o, \phi_o) &= f_r(\theta_i + \varepsilon, \phi_i, \theta_o + \varepsilon, \phi_o) \\ \forall \varepsilon, \omega &= (\theta, \phi) \end{aligned} \quad (5.5)$$

Other types of non-physical, analytic BRDFs are typically fast to compute and used in graphics where they produce satisfactory but approximate results for simple materials. This thesis is primarily interested in analytic, physical BRDFs introduced above, but does not enforce the requirements in experimentation.

Common models for BRDFs usually describe some form of diffuse reflection (view-independent), specular reflection (view-dependent highlights) or an additive combination of both. Linear separability is sometimes called Phong reflectance (distinct from the Phong BRDF discussed later).

$$f_r(\omega_i, \omega_o) = \rho_d f_d(\omega_i, \omega_o) + \rho_s f_s(\omega_i, \omega_o) \quad (5.6)$$

where ρ_d, ρ_s are the diffuse and specular albedos – wavelength dependent, scalar reflectivity values, which give rise to color¹⁸. Albedos are often written as part of the BRDF, and must be fractional values to obey energy conservation. $f_d(\omega_i, \omega_o)$ and $f_s(\omega_i, \omega_o)$ are diffuse and specular BRDFs, which individually satisfy the rules of physical BRDFs, if the total BRDF $f_r(\omega_i, \omega_o)$ is itself physical.

5.1.2 Methodology for Gonioreflectometry

Material BRDFs can be measured directly using a process called gonioreflectometry. Gonioreflectometry simply presents a methodical approach to sampling reflectance values across the hemisphere of lighting and view angles. A gonioreflectometer, in principle, requires no more than a light source, detector, sample material and a way to position these components. A reflectance ratio is collected along each of the four degrees of freedom (4-DOF) corresponding to the four dimensions of the BRDF – two illumination angles and two view angles (Figure 10). The reflectance data is then fitted to an analytic model or used (with interpolation) as a lookup table. This 4D sampling process is often quite slow, and more intelligent setups utilize automation and data parallelism with mirrors and cameras [Ward 1992; Marschner 1999] or reduce the dimensionality of the BRDF by assuming material isotropy [Marschner 1999]. Figure 10 illustrates a traditional 3-DOF gonioreflectometer setup compared with the static 3-DOF configuration, utilizing sparse data fitting developed for this thesis. This section will discuss motivations and justifications in this particular design.

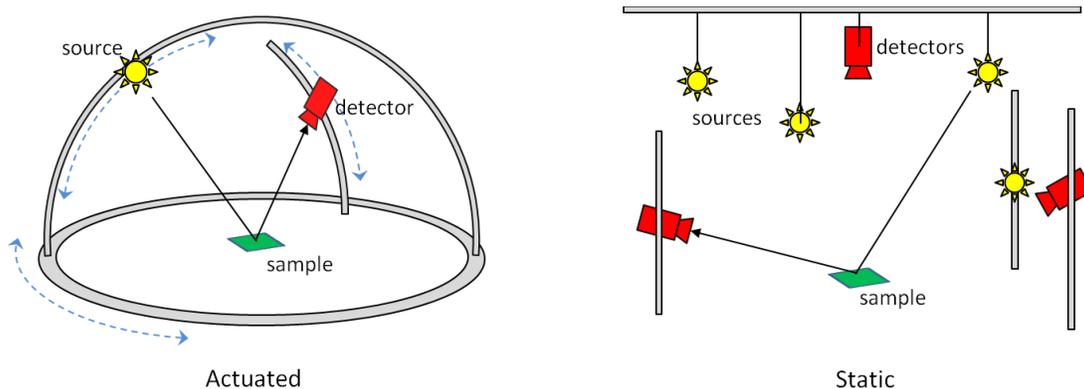


Figure 10. 3-DOF Gonioreflectometer Configurations. A traditional actuated setup where a single light source can move about the hemisphere of ray directions, but the detector is constrained to an arc with static azimuth (left). A redundant static

¹⁸ when observed with tristimulus detectors like the RGB filters of cameras or the cones of the human eye

setup developed in this work which utilizes multiple sources and detectors in the hemisphere; however, the cameras are coplanar (right).

5.1.2.1 Design

This work eschews the meticulous actuation of angular radiometry presented in prior work for a static, multi-camera based approach. An asymmetric but evenly distributed constellation of identical light sources and cameras instead capture the hemisphere of reflectance data from a centrally positioned, immobile sample. The sensors and sources are mounted in a “cube” configuration on a load-bearing aluminum frame. The major components are now described:

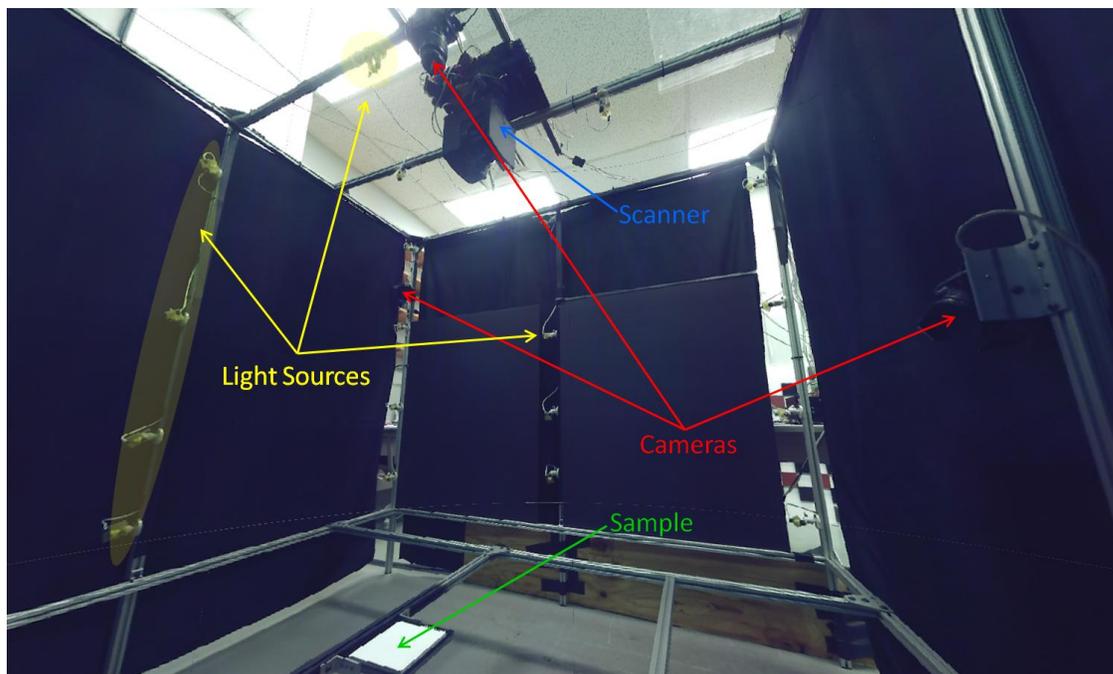


Figure 11. Photo of cube gonioreflectometer illustrating major components: fixed lights, cameras, test sample and LID AR for self-calibration.

Structure

Structural support for components is provided by a six foot cube aluminum frame, consisting of 80-20™ rails Figure 11. Attachment points are located on vertical side bars and horizontal bars across the top. One corner of the cube is utilized for mounting triangular supports which provide rigidity against twisting; components are therefore not located in this corner. The sample target (shown in Figure 11) is positioned at the center-bottom of the cube, elevated about a foot from the ground plane. The sample tray is located on movable rails to enable fine tune adjustment and shifting of the sample for non-standard views.

Errant light, interreflections and external illumination is a major concern for sensitive measurement. Consequently, several steps are taken to mitigate these effects. Firstly, all metals (such as aluminum supports) are covered in matte black gaffer tape to reduce reflections when possible. The faces of the cube are draped with cotton blackout doth to block external light (the lab is darkened prior to data acquisition, but secondary sources like monitors and electronic status lights remain). Lastly, sensors are either originally black or painted with a matte black spray. Optical glass, which cannot be covered for obvious reasons, from lensing is the only significant source of errant reflection in this setup.

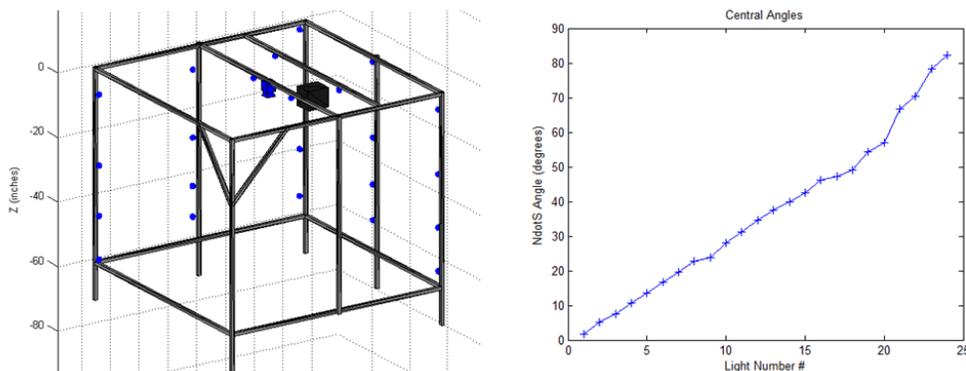


Figure 12. CAD model of cube frame with ideal light source positions (left). Plot of incident light rays from source positions distributed approximately uniform across the possible range of elevations (right).

Light Sources

Custom clusters of three tightly packed “super bright” 8mm LEDs are utilized as light sources (Figure 13 - center). Each LED draws 200mW of power (3.3V at 60mA) and radiates 11 lumens (a total of 33lm per cluster source) - about the brightness of a flashlight. Emitted light is a “cool white” color temperature of 5250K. Each LED is hemispherically diffused such that radiation is approximately isotropic; however the backplane is mounted to a swivel mount for fine tune adjustment. The entire cluster subtends an angular diameter of about 1 degree (fluctuating with distance).

A hemisphere of illumination is provided by 24 clusters, separated into “strings” of lights along the periphery of the cube. Vertical bars at the corners and midpoint of each face (and two on the top face) provide an equi-azimuthal distribution of strings. Lights are positioned such that the total distribution of incident light about the possible 90 degrees of elevation (Figure 12 - right) is almost uniform (when considering all lights).

Detectors

Three Canon EOS Rebel XS™ digital SLRs are utilized as the primary light detectors. These are 10 megapixel, 1.6x crop factor (22mm CCD) consumer cameras. Standard EFS 18-55mm lenses are configured at the 55mm, F/5.6 aperture setting. The cameras are mounted at 30, 60, and 90 degree elevations. A single Sick LMS-200 LIDAR, with rotating actuation to enable volumetric scanning, is mounted overhead for self-calibration of the sample location. It is additionally used for determination of macroscopic surface normals in non-planar samples and terrains.

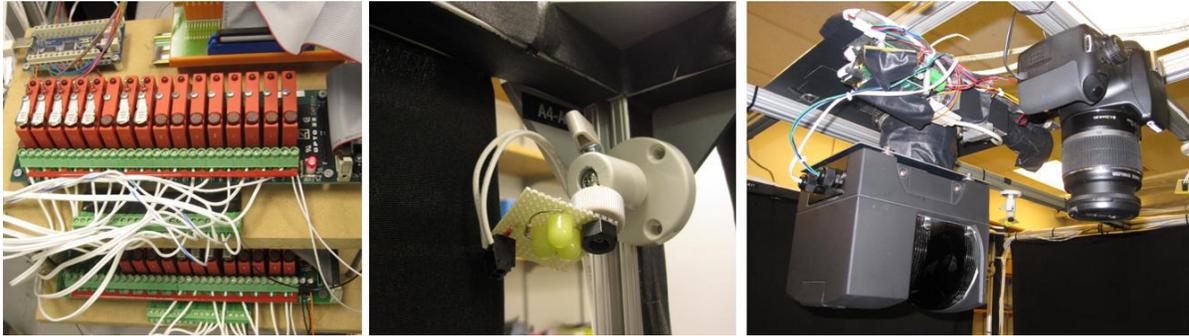


Figure 13. Detailed view of gonioreflectometer components. (Left) digital relays toggle power to light sources, (middle) each illumination source consists of a tight cluster of 3 LEDs each with their own hemispherical diffusers and (right) SLR camera and LIDAR mount positions at the top of the frame.

Electronics

Light sources are controlled from a PC via a USB digital I/O board. The board toggles a bank of relays which are able to drive the necessary higher operating currents (Figure 13 - left). The LIDAR scanner and actuating motor are driven over RS-232 serial. Lastly, data acquisition from the cameras also occurs over a USB connection. The entire setup is powered by two supplies at 5V and 24V.

Software

Control for the gonioreflectometer is entirely from within Matlab™ with C++ modules where necessary. Interface to the cameras is provided by the Canon SDK.

5.1.2.2 Discussion

There are several motivations for utilizing the described setup. Economy is perhaps the prevailing benefit; the cube frame is repurposed from a sensor characterization platform. All new components are off-the-shelf and readily available. While commercial spherical gantries may cost upwards of \$130,000 [stanford, Cammarano], this gonioreflectometer was developed for \$2,000. Secondly, static sensors and light sources greatly reduce device complexity. Construction and automation were simple as there are

no moving parts; and it was physically assembled in two weeks by a single person. This design was easily scaled to sizable proportions (2 meter cube) enabling larger-than-normal samples and parallel data acquisition. Lastly, the cube-like nature, while inefficient for gathering spherical data, is ideal for mounting additional sensors and secondary function as a solar simulator for artificial terrains (the use of which is presented later).

The cube gonioreflectometer does have several drawbacks. Static sensors present angular sparsity in measurement, particularly in view directions where each camera produces only one independent view point. Tiling the hemisphere with cameras would be prohibitively expensive, so only a single co-linear ring of cameras are utilized. These cameras share a relative azimuth but have differing elevations, giving the sensor three true degrees of freedom. Overall view sparsity is mitigated to some extent by utilizing large samples and the conical nature of measurement from cameras. This approach requires large, flat and pure samples which are hard to obtain for many materials. Consequently, powdered samples, which can be manipulated in a flat tray, are utilized when possible. This does change the reflective nature of the material somewhat; however, many of these materials are found naturally in particulate form in the domain. Ultimately a data fitting approach is utilized instead of a reflectance lookup table. This approach compliments relative data sparsity and to reduces noise.

The experimental setup explored here does not replace traditional gonioreflectometers, as the underlying purpose (and acceptable quality) of the data utilized in this thesis are different. This approach, in fact supplements traditional gantry methods in field simulation and experimentation. It is a cheap, fast, and low data burden sensor that enables commodity radiometry with oversize samples. Notably, some materials analyzed in this work have been spectroscopically analyzed with precision instrumentation - this includes JSC-1A and likely others [Cord, et al 2003; Johnson, et al 2008; Mustard, et al. 1989]. In a strict theoretical sense, it is possible to derive or infer visible-spectrum reflectance models from this existing data. However, the purpose of prior experimentation is materials science not robotic perception. Resulting analysis neither considers fitting for common perceptual BRDFs which can be inverted for computer vision, nor aggregate appearance as measured by vision sensors.

The application intent of this thesis is to utilize robots with commodity RGB cameras in the visible spectrum. These cameras have peculiarities of lensing, CCD circuitry, radiometric curves, and discrete spectral sensing. There is no better validation of this than imaging these materials with those very cameras in a controlled setting.

5.1.2.3 Calibration

Calibration is required to reduce errors and increase certainty in the data. It is simply a systematic approach to recovering unknown parameters necessary for radiometry: geometry, camera response, and radiative transfer. The procedure for this gonioreflectometer is particularly involved, necessitating three interconnected optimizations. This stems from the rectangular nature of the frame, which enforces radially varying mounting distances for the cameras and light sources. In fact, an additional novel calibration for incident irradiance is required as compared to traditional setups.

Geometric Calibration

The “as-built” geometry of the components requires great care in calibration, particularly as compared to a spherical gantry, where components are favorably affixed to an arm of constant radius. These positional and directional uncertainties introduce additional sources of *vector* measurement error not present in spherical configurations. In lieu of a traditional photometric approach based on optimizing camera reprojection error in visible sources, a direct measurement of positions is performed utilizing a Faro Focus3D survey LIDAR. This LIDAR has 1mm accuracy and sufficient angular resolution to identify the components in the resulting point cloud [Wong, et al. 2011]. As such, the accuracy of this direct method is likely an order of magnitude greater than any camera based approach. The major source of error in geometric calibration are primarily the uncertainty between detection of a component’s position as the centroid of a discrete point “blob” and the true mounting position. Human validation is used to minimize the possibility of errant matches.

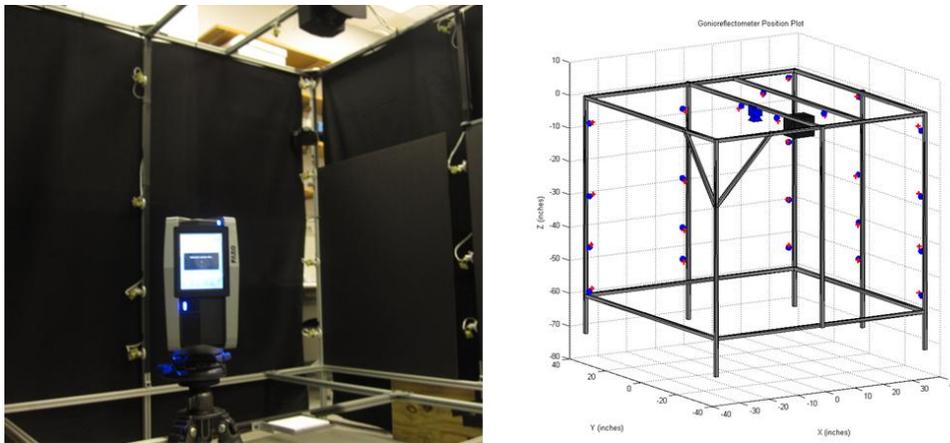


Figure 14. Geometric calibration utilizing an “as-built” survey scanner (left) and model of cube showing planned (blue) and as-built (red) positions of the light sources (right).

Figure 14 illustrates the as-built positions of the light sources compared to the intended positions. The cube's coordinate frame is given as x-right, y-back and z-up (from the entry point of the lab) with the origin at the top-center where the 90 degree camera is mounted.

Camera Calibration

Camera calibration – which involves separate radiometric and geometric components – is also required for each camera. Geometric camera calibration recovers the optical transfer function of each lens, which maps every pixel to a unique incoming light ray. This enables use of the sample as a region of measurements in contrast to a single mean direction from the target centroid to the camera optical center. Calibrations are performed with the Matlab camera calibration toolbox [Bouquet 2001]. Lens field of view, focus and aperture are locked during and after calibration as they are unique to each camera position.

Radiometric calibration recovers the response curve of the camera. This function is a mapping from pixel value to irradiance incident on the sensor - an effective inverse of the “gamma” function and other artifacts of analog-digital conversion. The function enables transformation from nonlinear units of pixel value - usually uint8 values - to linear (but relative) units in the range $[0, \infty]$. With additional knowledge of exposure settings such as aperture, shutter speed and gain (ISO) at the time of image capture, physical radiance units can be recovered. These are ultimately needed for accurate reflectance measurement. One mapping exists per color channel, therefore the total radiometric function is (RGB) vector valued.

The radiometric curve is recovered by imaging a static scene multiple times while modulating shutter speed – a known linear quantity in image formation. Pixels in the same location across images are only different in exposure and form related diques. If all observed pixel values sufficiently span the space of possibilities, a polynomial function can be fit to describe this relationship. This process is described in detail in [Debevec, et al. 1997].

Radiative Compensation

It is well known that radiant intensity of small sources falls off (light attenuation) over distance. Moreover, it cannot be assumed that all light sources are spectrally and radiatively identical due to uncertainty in manufacturing, wiring resistance and coating. A calibration procedure must be conducted to normalize the incident irradiance from each light source and camera combination. This calibration can only be accomplished after calibration of positions and camera responses.

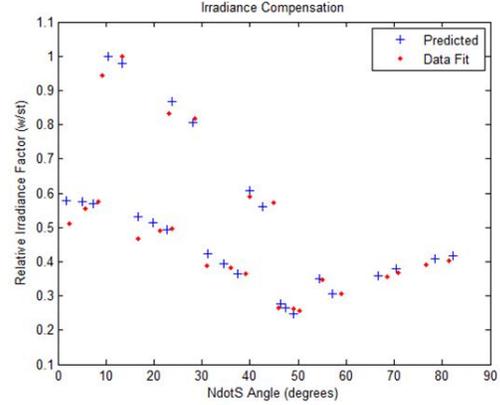
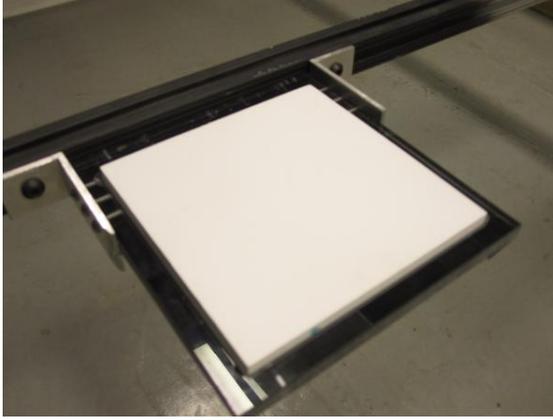


Figure 15. Calibration of incident irradiance using an ideal Spectralon sample (left) and compensation for distance falloff from light source positions on cube frame (right). The sources are small enough that the incident light on the target does not deviate significantly from ideal.

The approach to normalizing irradiance involves comparing data of a reference target to predicted ideal values and fitting a function from source and camera distance to measured intensity. The target used in calibration is made of Spectralon, which is 99% reflective and is assumed to be a perfect Lambertian reflector. An ideal point source is known to fall as the square of distance; however this assumption is rarely valid as even the smallest sources subtend a nontrivial area. Instead, a polynomial of the following form is fit:

$$\begin{bmatrix} S_1 \\ \dots \end{bmatrix} = \begin{bmatrix} \frac{1}{d_1^2} & \frac{1}{d_1} \\ \dots & \dots \end{bmatrix} \times \begin{bmatrix} A \\ B \end{bmatrix} \quad (5.7)$$

$$x = S^{-1}d$$

where d_i is the distance of the i -th source or view combination to the target and S_i is the irradiance compensation factor. The coefficients of the polynomial are found with least squares regression. This polynomial is used in lieu of direct lookup to prevent overfitting.

In this calibration step, whitebalancing is also performed for each of the lightsources utilizing knowledge of the spectral response of the Spectralon. Whitebalancing simply finds scalars $\hat{w} = \langle w_r, w_g, w_b \rangle$ such that $w_r \times E_{i,r} = w_g \times E_{i,g} = w_b \times E_{i,b}$ where $E_{i,r}$ is the radiance in the red channel of the i th light source and so on.

5.1.2.4 Data Acquisition

Data acquisition is performed by illuminating each of the 24 sources consecutively, one at a time, and capturing the sample with each camera (which can be parallelized). This means a total of 72 unique lighting and view combinations are captured. Multi-pixel acquisition of the target area from each camera boosts this to a total 1152 illumination-view combinations. Cameras take a bracket of LDR images at [0.5, 1, 2] seconds. HDR images are created in post process to reduce noise and expand the range of detectable reflectances (component LDR images are 256 bit TIFFs). Total acquisition takes about 10 minutes per sample. A human manually selects the extent of the sample in each of the three principal views as the sample may not completely cover the target area. Sixteen intersecting rays in the target area are produced and intensity data is averaged to those sample locations. Finally, raw pixel values are transformed to irradiance utilizing calibration constants and compensated for distance disparities in each of the sources.

5.1.3 Data-Fitting Analytic BRDFs

This section describes the BRDFs of interest and the approach to data-fitting. A computer graphics vector notation for radiant transfer is now introduced. These formulations utilize world space instead of surface-relative spherical coordinates, and more directly reflects the geometric nature of the gonioreflectometer and vision in general. It is possible to convert between these coordinate systems with simple vector math.

Table 2. Static Parameters common across BRDFs

Symbol	Description
\vec{L}, \hat{L}	Light source (incoming) intensity and direction and normalized direction vectors, respectively
\hat{N}	Surface normal vector
\hat{V}	Viewing (outgoing) direction vector
\hat{H}	Half angle (between source and viewer) vector

Common parameters across all BRDFs are listed in Table 2; these are integral in the calculation of BRDFs but do not distinguish between them. These values are considered static in optimization and are given by virtue of calibrated measurement in the gonioreflectometer. Each parameter is a geometric vector with normalized (unit magnitude, \hat{x}) form, while only the unnormalized form (\vec{x}) of the L vector is significant. These vectors are unique for each scene point.

The light source direction (\hat{L}, \vec{L}) is a vector from each scene point to the center of the illuminant (considered to be an ideal point source). The unnormalized light vector measures both the direction and source intensity, I_s , in lumens. This is distinct from prior use of the *illumination* vector in context of gonireflectometer calibration, which measures a distance. In some related work, only the normalized vector is used and there is a separate scalar for source intensity is utilized; these forms are semantically equivalent ($\vec{L} = \hat{L} \times I_s$). \hat{N} is the surface normal vector. As every sample material is a flat surface, this is simply the vector $\langle 0, 0, 1 \rangle$. \hat{V} is the view (also known as eye) vector from each scene point to a particular camera's optical center. Finally, \hat{H} is the half angle vector which bisects the view and source direction ($\hat{H} = \frac{\hat{L} + \hat{V}}{\|\hat{L} + \hat{V}\|}$). This is utilized to determine the deviation from the ideal mirror reflection direction.

The radiance in view direction (\hat{V}), from a linear combination of infinitesimally small point sources is given by the modified rendering equation of section 2.3. This equation, rewritten with the above graphics terminology becomes:

$$L_o(\hat{V}) = \sum_i f_r(\hat{V}, \hat{N}, \hat{L}) (\hat{N} \cdot \hat{L}_i) I_i \quad (5.8)$$

This form is of particular interest due to simplicity in model fitting and rendering – in which it is used heavily in this thesis. Notice that it is a product of the BRDF (f_r) and a cosine projected irradiance from each source, $(\hat{N} \cdot \hat{L}) I = \hat{N} \cdot \vec{L}$. This term, called Lambert's law, is typically written as part of the BRDF for completeness in graphics literature.

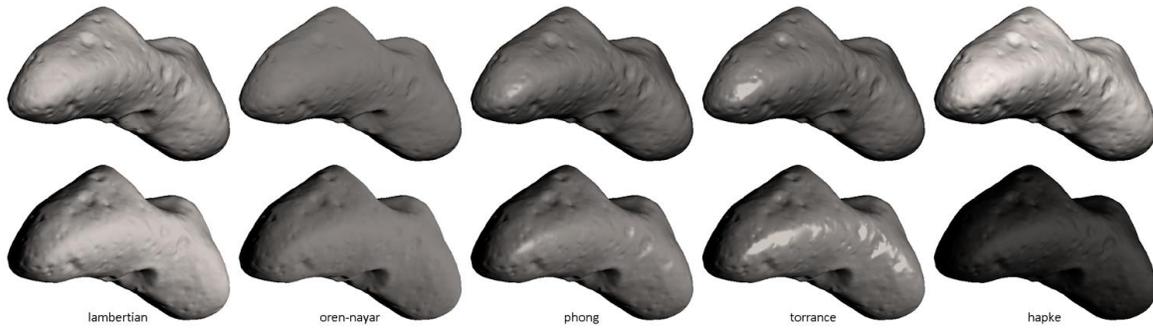


Figure 16. Canonical illustrations of the five BRDFs analyzed: (left to right) Lambertian, Oren-Nayar, Phong, Cook-Torrance-Sparrow and Hapke. They are rendered on the Eros model with straight on (top row) and glancing (bottom row) illuminant directions.

Five analytical BRDFs are utilized in fitting reflectance data. In ascending order of number of fitting parameters, they are: (1) Lambertian, an ideal diffuse model; (2) Oren-Nayar, a rough microfacet diffuse model; (3) Phong, an empirical smooth specular model; (4) Cook-Torrance-Sparrow, for smooth specular microfacet reflections, and (5) the Hapke model for backscattering materials. Table 3 summarizes these models and their canonical appearances are illustrated in Figure 16. Descriptively, Lambertian and Oren-Nayar materials appear matte and similar when viewed from different angles, while the other Phong, Torrance and Hapke Models exhibit a variety of “shininess” effects in certain key directions.

Table 3. Analytical BRDFs utilized in Reflectance Analysis

BRDF	Description	# Variables
Lambertian	Ideal diffuse	1
Oren-Nayar	Rough, microfacet, diffuse	2
Phong	Smooth specular	3
Cook-Torrance-Sparrow	Smooth, microfacet specular	4
Hapke/Lommel-Seeliger	Multiple scatter, diffuse	5

These BRDFs were selected such that increasing dimensionality of parameters coincides with an increasing capability of representing complex phenomena. However, most models advantageously reduce to Lambertian with appropriate zeroes. Thus, materials that are an ideal subset of two or more models should produce equivalent responses in related BRDFs. Though this is not always strictly true in slope climbing optimization due to selection of the fitness function and input noise, the degree of relatedness can be easily surmised from empirical data. This fact can be exploited to ensure a “simplest

explanation” coincident with the principle of Occam’s razor. As this thesis concerns mostly diffuse materials and the representation of appearance as Lambertian, this reducibility is of key importance in analysis.

Some common analytic BRDFs are not considered due to scope and applicability. The Dirac mirror function (a perfectly specular reflectance) is impractical to fit - only a single view direction produces a response - and virtually nonexistent in the domain of interest. Anisotropic models like Ashikhmin-Shirley [Ashikhmin, et al. 2000] and Ward [Ward 1992] cannot be tested as there is insufficient sampling of viewpoints. This is an artifact of the repurposed gonireflectometry setup presented this work. Cameras are mounted in a numerically-coplanar ring of elevation angles and are only capable of sampling three of the four dimensions in unconstrained surface reflectance. Reflectance data collected this way can only discriminate BRDFs modulo rotational (azimuthal) uncertainty. The complexity offered by this extra degree of freedom explains mostly “brushed metallic” phenomena due to a bias in banding or orientation of surface grains in a material. Fortunately, these are rarely encountered in the domain of interest. The vast majority of materials – particularly in planetary settings – are isotropic.

Several BRDFs are modified from their original forms utilizing common approximations for complex terms, principally to facilitate data fitting. These modifications are known to violate strict radiosity in some cases. Four of the models presented here are based on the physical principles of geometric optics, with the Oren-Nayar and Torrance models utilizing approximations. The fifth – Phong reflectance – is strictly empirical in nature. While this thesis strives to identify and utilize physics-based explanations of appearance, strict, energy-conserving radiative transfer is not a primary purpose for this analysis. Physical BRDFs of materials can be acquired with greater accuracy and density using any of the dedicated gonireflectometry setups described in prior work.

The ease of use, degree of fit, applicability to computer vision algorithms as well as physical plausibility of BRDFs are all equally considered here. The Phong model, for example, is often a good approximation that has the mathematical advantages of simplicity. Consequently, a large collection of existing computer vision algorithms operate on Phong-type materials [Vogel, et al. 2009]¹⁹. The idea is that this work can determine the degree to which a Phong - or other - assumption can explain the appearance of planetary materials, enabling use of these techniques.

¹⁹ Phong reflection has also been utilized successfully in solving physically-based radiative transfer anomalies on the NASA Pioneer spacecraft [Francisco, et al. 2011].

The following section overviews these BRDFs, their unique parameters, and implementation specifics in the context of this thesis. The reader is encouraged to review the primary sources for further detail regarding derivations and physical explanations. The terminology introduced here is mostly adapted from these sources, with some minor modifications to enhance clarity and consistency between models.

5.1.3.1 Lambertian

Parameter	Description	Range
ρ_d	Diffuse albedo	[0,1]

The Lambertian BRDF is the simplest model and encodes a perfectly diffuse reflectance. It consists of a constant scalar albedo, ρ_d , and Lambert’s law.

$$f_{Lambertian}(\rho_d) = \rho_d (\hat{N} \cdot \vec{L}) \quad (5.9)$$

Most notably Lambertian materials do not depend on viewing angle – they appear the same from any direction unique only to a combination of surface normal and source vectors. This effect is caused by subsurface body interactions, which scatter light isotropically, and gives rise to the intrinsic color of the material [Zickler SUV paper]. Lambertian materials are of primary importance in this work as they enable a vast collection of computer vision techniques.

The albedo is the only fitting parameter for this BRDF. It is noted that this function is often written with a $\frac{1}{\pi}$ normalizer to account for conservation of energy across the hemisphere (assuming unitary radiance from the light source). However, the true radiance of the light source and the metric units of the detector (camera) are not known. Only relative comparisons can be made with a reference material that is assumed to be perfectly reflecting. This and other scalar normalizations are therefore factored into a preliminary data transformation before optimization. These normalizers are thus omitted for brevity in further BRDF discussion such that albedos are scaled in the interval [0,1].

5.1.3.2 Oren Nayar

Parameter	Description	Range
ρ_d	Diffuse albedo	[0,1]
σ	Roughness factor (standard deviation of microfacet angles) in radians	$[0, \frac{\pi}{2}]$

The Oren-Nayar model presented in [Oren, et al. 1993] is a generalization of Lambertian reflectance for rough surfaces. Surfaces are assumed to be collections of randomly oriented *microfacets* which are paired in symmetric v-shaped cavities. While each facet is too small to be individually considered, the macroscopic distribution (and its radiance transfer) is estimable and assumed to be a zero-mean Gaussian. The facets are assumed to be much larger than the wavelength of light, such that radiosity can be analyzed with the rules of geometric optics. The projected radiance of an ideal lambertian facet can then be calculated.

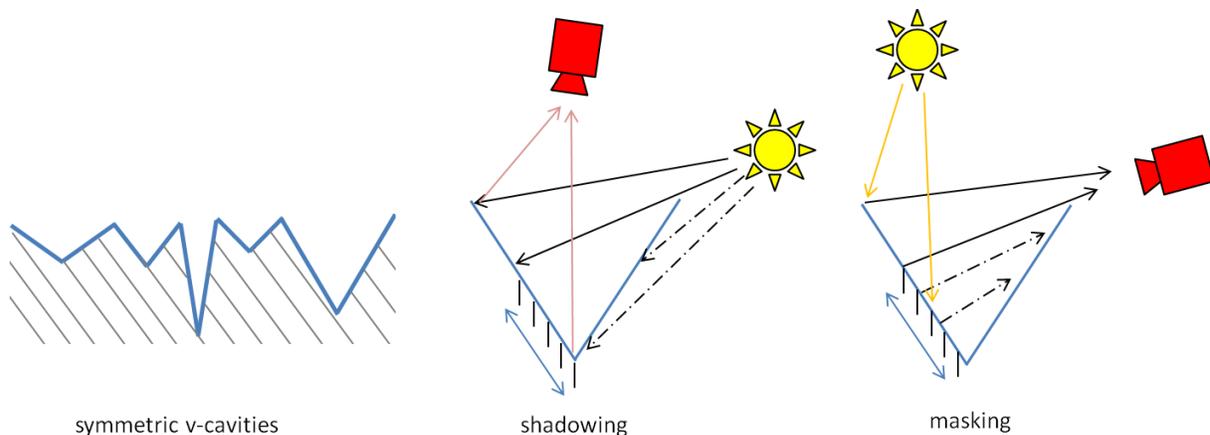


Figure 17. Microfacet Surface model utilized in Oren Nayar and Torrance BRDF models. Surfaces are randomly oriented collections of microscopic facets arranged in symmetric v-shaped cavities (left). Light reflecting from the source to the viewer is attenuated by the surface geometry by shadowing, where the source is blocked (middle) and masking, where the viewer is blocked (right).

This radiance is attenuated by *self shadowing* and *light masking* – summarized by a fractional product known as the geometric attenuation factor (GAF) – and increased by additive inter-reflections in each cavity (see definition of Torrance model in the following section for more information). The contribution of each facet is finally integrated across all the normals according to the distribution to find the total radiance emitted from a scene patch. The final formula is a mixture of analytical approximation and numerical integration as the resulting integral is not easily evaluated (5.10). The meanings of the individual components are not particularly intuitive.

$$f_{\text{Oren-Nayar}}(\rho_d, \sigma) = \rho_d (\hat{N} \cdot \bar{L}) \times (\zeta + \chi + \varepsilon) \quad (5.10)$$

where,

$$\begin{aligned} \zeta &= 1 - 0.5 \times \frac{\sigma^2}{\sigma^2 + 0.33} \\ \chi &= \gamma \times A \times \tan \beta \\ \varepsilon &= (1 - |\gamma|) \times B \times \tan\left(\frac{\alpha + \beta}{2}\right) \\ A &= \begin{cases} 0.45 \times \frac{\sigma^2}{\sigma^2 + 0.09} \sin \alpha & \gamma \geq 0 \\ 0.45 \times \frac{\sigma^2}{\sigma^2 + 0.09} \left(\sin \alpha - \left(\frac{2\beta}{\pi}\right)^3 \right) & \gamma < 0 \end{cases} \\ B &= \frac{1}{8} \left(\frac{\sigma^2}{\sigma^2 + 0.09} \right) \times \left(\frac{4\alpha\beta}{\pi^2} \right)^2 \end{aligned}$$

and angle aliases are given by,

$$\begin{aligned} \alpha &= \max \left[\text{acos}(\hat{V} \cdot \hat{N}), \text{acos}(\hat{L} \cdot \hat{N}) \right] \\ \beta &= \min \left[\text{acos}(\hat{V} \cdot \hat{N}), \text{acos}(\hat{L} \cdot \hat{N}) \right] \\ \gamma &= \left\langle \hat{V} - \hat{N}(\hat{V} \cdot \hat{N}), \hat{L} - \hat{N}(\hat{L} \cdot \hat{N}) \right\rangle \end{aligned} \quad (5.11)$$

In the original paper, the diffuse interreflection term (not shown above) is discarded as it contributes only minimally to the total radiance while drastically increasing complexity and decreasing quality of data fitting. It is omitted in the work presented here as well. While the vector calculations are complex, there are only two unique parameters to the model - the diffuse albedo (ρ_s) and the standard deviation (σ) of the the microfacet angles. It is notable that when $\sigma = 0$ (no slope deviations) the Oren-Nayar model reduces to the Lambertian model.

5.1.3.3 Phong

Parameter	Description	Range
ρ_d	Diffuse albedo	[0,1]
ρ_s	Specular albedo	[0,1]
α	Specular hardness	[0,∞]

While ideal specular reflection occurs at a single “spike” viewpoint, realistic materials often exhibit less jarring highlights. Phong reflection is a phenomenological model that approximates a smooth specular falloff at views near the mirror direction [Phong 1975]. Total radiance is a combination of diffuse and specular effects, where the diffuse component is the Lambertian model. The specular component is an exponent of the cosine relationship between the ideal reflection (R) and view directions (v). The degree of this exponent controls the “hardness” of the specular lobe. Small exponents describe broad, low-gradient highlights while large exponents increasingly describe a mirror spike. The diffuse (ρ_d) and specular (ρ_s) albedos along with the specular exponent (α) comprise the fitting parameters for the Phong model.

$$f_{Phong}(\rho_d, \rho_s, \alpha) = \rho_d (\hat{N} \cdot \bar{L}) + \rho_s (\bar{R} \cdot \hat{V})^\alpha \quad (5.12)$$

where,

$$\bar{R} = 2(\hat{N} \cdot \bar{L})\hat{N} - \bar{L}$$

is the reflection vector. An additional constraint is that the specular and diffuse albedos must sum to less than 1 ($\rho_d + \rho_s \leq 1$) so that the material is not self-emitting; this is observed in optimization. This constraint alone is not sufficient to satisfy energy conservation and the model is known to “leak light”. However, the succinctness and linear form of the function is of great interest in model fitting and inverting for vision application.

It is noted that in the original formulation, Phong suggests an *ambient* offset radiance term to account for interreflections. This term strongly violates energy conservation and it is not utilized in this work. The model reduces to Lambertian with zero specular albedo.

5.1.3.4 Cook-Torrance-Sparrow

Parameter	Description	Range
ρ_d	Diffuse albedo	[0,1]
ρ_s	Specular albedo	[0,1]
r	Mean microfacet slope	[0,∞]
F_0	Initial value of the Fresnel function at normal incidence, as used in Schlick's Approximation	[0,1]

The Cook-Torrance-Sparrow model (also known as Cook-Torrance and Torrance-Sparrow) is a microfacet surface model on which the Oren-Nayar BRDF is based. However, each of the facets here is assumed to be a perfectly reflecting mirror. Thus, only mirrors oriented in the half-way direction contribute to the total radiance of the surface patch. The BRDF is linearly separable into diffuse and specular terms much like the Phong model with the diffuse term being the Lambertian BRDF [cook torrance 1982]. The specular reflection is a physically-plausible combination of three factors: Fresnel reflectance (F), surface roughness (D) and geometric attenuation (G).

The Fresnel reflectance (F) describes an angular dependency between the incident illumination and the magnitude reflected from a surface. In grossly simplifying terms, this relationship states that the intensity of specular highlights from grazing angles is greater than those for normal incidents. Intuitively, this effect is governed by the indices of refraction for the typically air-material interface. In normal incidents, greater energy is transmitted through the body of the material, while in glancing incidents most energy is reflected. The Fresnel term is distinct from – and acts in opposition to - Lambert's law which describes diminishing intensity from increasing areas of equivalent flux at glancing angles. As index of refraction is wavelength dependant, so is the Fresnel term; which leads to color shifting in specularities.

The surface roughness term (D) controls the bandwidth of the specularity; it is a distribution of facet slopes that describes the percent of facets oriented in the mirror direction. Rough surfaces produce highly directional specularities, while smooth surfaces reflect broadly. Several possibilities for slope distributions are given in Cook's paper including an approximating asymmetric Gaussian and the physically plausible Beckman distribution, derived from microfacet theory.

The geometric attenuation factor (G , also utilized Oren-Nayar) encodes the reduction of light reaching the viewer due to interactions with facet geometry. This factor is thus a fractional scalar relative to a

perfectly transmitting facet of the same geometry. Light is attenuated by two similar processes, shadowing and masking (Figure 17). Shadowing is the partial blocking of incoming light that would normally reach the surface due to an occluding facet, while masking is the blocking of outgoing light. The GAF can be succinctly represented as a minimum of two vector dot product tests for these conditions.

The full Cook-Torrance-Sparrow BRDF is given in equation(5.13).

$$f_{Torrance-Sparrow}(\rho_d, \rho_s, r, F_0) = \rho_s (\hat{N} \cdot \vec{L}) \times \frac{F \times D \times G}{(\hat{N} \cdot \hat{V}) \times (\hat{N} \cdot \hat{L})} + \rho_d (\hat{N} \cdot \vec{L}) \quad (5.13)$$

where,

$$F = F_0 + \left(1 - (\hat{H} \cdot \hat{V})\right)^5 \times (1 - F_0)$$

$$D = \frac{1}{\pi r^2 \times (\hat{N} \cdot \hat{H})^4} \times \exp \left[\frac{(\hat{N} \cdot \hat{H})^2 - 1}{r^2 \times (\hat{N} \cdot \hat{H})^2} \right]$$

$$G = \min \left[1, \frac{2(\hat{N} \cdot \hat{H})(\hat{N} \cdot \hat{V})}{\hat{V} \cdot \hat{H}}, \frac{2(\hat{N} \cdot \hat{H})(\hat{N} \cdot \hat{L})}{\hat{V} \cdot \hat{H}} \right]$$

F is a Fresnel term approximation, D is the Beckman Distribution for surface roughness, and G is the Geometric Attenuation Factor. In this work, fitting the Torrance model comprises optimization over four parameters: diffuse albedo (ρ_d), specular albedo (ρ_s), root-mean-squared slope of the facets (r), and a fresnel term initial value (F_0). There is an additional constraint (similar to the Phong model) that $P_d + P_s \leq 1$.

This form of the Torrance model deviates from Cook's paper in the formulation of the Fresnel term, which is replaced with a functional form widely used in graphics called Schlick's Approximation [Schlick 1994]. Use of this approximation is two-fold. Firstly the original Fresnel equations are wavelength dependent; complex spectral functions of indices of refraction can neither be resolved by the experimental setup nor fit to empirical data with any certainty or reasonable values. Secondly, the function distinguishes between polarized and unpolarized light, and the degree to which each comprises the total illumination. Polarization is likewise not detected by this setup. Schlick's Approximation is wavelength invariant and polarization invariant (though, materials can consist of multiple responses in

the color spectrum, i.e. RGB values). It requires only a single parameter (F_0) which is the initial value of the Fresnel function at unpolarized, normal incidence at the dominant wavelength. F_0 can be calculated from a known index of refraction of the material, or estimated directly as done here. A table of Fresnel coefficients for common materials, interfaced with a vacuum, is reproduced from [Real-Time Rendering 3rd Edition] for comparative purposes. This work assumes non-metallic Fresnel reflectance, meaning that that F_0 is constant across color channels.

Table 4. Table of Common Fresnel Coefficients

Insulator	sRGB F_0	Conductor	sRGB F_0
Water	[0.15, 0.15, 0.15]	Gold	[1.00, 0.86, 0.57]
Glass	[0.21, 0.21, 0.21]	Silver	[0.98, 0.97, 0.95]
Plastic	[0.24, 0.24, 0.24]	Copper	[0.98, 0.82, 0.76]
Ruby	[0.31, 0.31, 0.31]	Iron	[0.77, 0.78, 0.78]
Diamond	[0.45, 0.45, 0.45]	Aluminum	[0.96, 0.96, 0.97]

*reproduced from Real-Time Rendering 3rd Edition

The Beckman distribution is utilized as the surface roughness term. This distribution requires a single parameter r , the root-mean-squared slope of the facets; small values of r give a smooth surface with highly directional specularity, while large values give a broad specularity.

The Torrance model reduces to Lambertian with zero specular albedo or if the mean slope parameter tends to infinity. The latter case is technically permissible, but highly unlikely in optimization.

5.1.3.5 Hapke Lommel-Seeliger

Parameter	Description	Range
w	Single scattering albedo	[0,1]
b	Particle phase parameter 1	[0,1]
c	Particle phase parameter 2, the fraction of energy which is backscattering	[0,1]
B_0	Magnitude of opposition surge	[0,1]
h	Width of opposition surge	[0,1]

The Hapke model is used in remote sensing to explain the perceived brightness of regolith-covered planetary bodies. It primarily explains the opposition effect - a greater than predicted radiance when the viewer and source are nearly co-aligned. Phenomena explained by this model include the bright “halos” on the ground (*heiligschein*) in Apollo lunar images where the sun is behind the camera and the limb effect of the moon. As such, this is an interesting BRDF for the planetary environments explored in this thesis. However, due to the large number of parameters, and questionable applicability at the macroscopic (robot-sized) scale, the Hapke model is notoriously difficult to invert [Liang, et al. 1996].

The principal physical explanation for the opposition effect is shadow hiding, where particles mask or shadow light paths except when the illuminant and viewer are colinear. This is not unlike the model assumed in the Oren Nayar BRDF. Output radiance is defined as a sum of single scattering (which can be forward or backscattering in contrast to Oren-Nayar) and isotropic (equally in all directions) multiple scattering of light rays in the body of the material in addition to a direct component. The magnitude of single scattering is modulated by the opposition surge.

The Hapke model utilized here is given as:

$$f_{Hapke-LS}(\rho_a, \omega, b, c, B_0, h) = (\hat{N} \cdot \hat{L}) \times \frac{w}{4\pi(\mu_0 + \mu)} \times ((1+B)P + H(\mu_0)H(\mu) - 1) \quad (5.14)$$

where,

$$B = \frac{B_0}{1 + \frac{1}{h} \times \tan\left(\frac{g}{2}\right)}$$

$$P = (1-c) \times \left(\frac{1-b^2}{(1+2b \times \cos g + b^2)^{3/2}} \right) + c \times \left(\frac{1-b^2}{(1-2b \times \cos g + b^2)^{3/2}} \right)$$

$$H(x) = \frac{1+2x}{1+2x(1-w)^{1/2}}, \quad x = \{\mu, \mu_0\}$$

where angle and foreshortening aliases are given as:

$$\mu_0 = \hat{N} \cdot \hat{L}$$

$$\mu = |\hat{N} \cdot \hat{V}|$$

$$g = \text{acos}(\hat{N} \cdot \hat{V})$$

An important alias utilized here is g , which defines the phase angle - the interior angle between the view and source directions. The principal components are: $B(g)$, the shadow-hiding opposition effect; $P(g)$, the Henyey-Greenstein particle phase function which describes the angular distribution of single scattering, and $H(x)$, the Chandrasekhar function for isotropic multiple scattering. The final radiance is modulated by $\frac{w}{4\pi(\mu_0 + \mu)}$, the Lommel-Seeliger coefficient, which determines the magnitude of scattering when the surface roughness is much larger than the wavelength of light.

The classical Hapke model comprises 6 unique parameters, known as the Hapke parameters. The parameter w is the single scattering albedo: the albedo of small particles which interact with the light. This is distinct from the *bulk* albedo utilized in prior BRDFs and a particulate material will take on different values for each. Two particle phase parameters, b and c , are material properties which are first and second order coefficients of the Legendre polynomial approximation to the nature of scattering. Qualitatively, b determines the width of the scattering lobe (large b is narrow), and c is a fraction determining whether scattering is primarily forward ($c < 0.5$) or backward ($c > 0.5$). B_0 is the magnitude of the opposition effect, which is typically near 1 for most regolith [Pugacheva, et al. 2005]. The density, porosity and compaction of the media is described by h , which modulates the angular broadness of the opposition effect [Helfenstein, et al. 1987]. Finally, a parameter θ measures the macroscopic roughness of microfacets in the terrain, akin to those in Oren-Nayar or Torrance. The original formulation of the Hapke model enabled use of a custom BRDF with input θ to further modulate the radiance based on macroscopic effects. This parameter is not utilized in this work as it assumes knowledge of the intrinsic BRDF (and further assumption of microfacets) of the material which is a self-defeating exercise. Instead, a macroscopic Lambertian BRDF is assumed which does not utilize θ and hence has no effect in optimization.

This model does not reduce readily to the Lambertian model. Furthermore, it is well known that when an inaccurate Hapke model is used for inversion, the retrieved parameters may have no physical analogues [Shepard, et al. 2007]. For example, determination of the asymmetry parameter from measurements is ill conditioned [Liang, et al. 1996]. Thus, steps must be taken in optimization to physically constrain the space of output variables. Here, estimates of the albedo from a Lambertian optimization are used as strict bounds (in the interval (66% to 133%)) for the single scattering albedo.

Furthermore, the height of the opposition effect B_0 is regularized to be near 1 as estimated in prevailing literature.

5.1.4 Experimental Results for Materials

This section presents BRDF fitting results for several materials of interest in planetary and subterranean domains as well as arguments for the significance of these materials. A total of nine materials were analyzed: **two** reference materials, Spectralon™ and matte paint; **two** materials from a planetary (lunar) environment, JSC-1A regolith simulant and CMU-1 simulant; and **five** materials from terrestrial underground environments, coal dust, limestone dust, spray concrete (gunite), granite, and black sandstone. Furthermore, the macroscopic BRDF of CMU-1 covered lunar terrain (featuring rock-sized features with surface roughness measured with a LIDAR scanner) was recovered, but those results are presented in a future section.

The five analytical BRDFs presented in the prior section are each fitted to the observed radiance from the camera sensor by minimizing squared error between model-predicted radiance and observed radiance. As individual color pixels are not truly independent (due to Bayering) and the camera exhibits moderate noise in low light, fitting is not performed for each channel individually. Instead, white balancing is performed on each image which is then transformed to XYZ tristimulus values. The Y channel (luminosity) is utilized as the measured radiance value. Color data is independently averaged across all measurements from the white balanced images and stored separately for rendering.

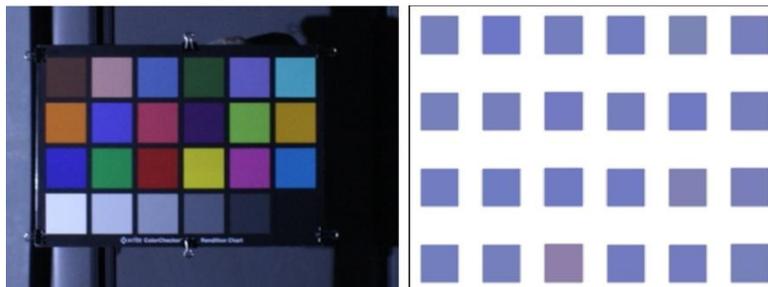


Figure 18. Macbeth Chart (left) used for color calibration and estimated whitepoint for each light source (right).

Optimization over input parameters is performed using a bounded simplex search with linear constraints in the form $Ax \leq c$. This type of numerical optimization, while highly efficient, is fairly susceptible to being trapped in local minima. To prevent this, some hand tuning is utilized to maintain sensible results.

Estimated albedo from a much more robust Lambertian fitting is used as the initial albedo value for higher dimensional functions. Moreover, for the Hapke function, maximum bounds are calculated from the Lambertian albedo to prevent the function from numerical instability.

Two independent metrics are used to determine the quality of fit and also the “best” BRDF for a material. The primary metric, root-mean-squared (RMS) error ε is a derivative of the objective function used in optimization.

$$\varepsilon_{RMS}(x) = \sqrt{\frac{1}{n} \sum (E_i - I_i)^2} \quad (5.15)$$

where E_i is the observed radiance at sample i and $I_i = f(x, n_i, l_i, v_i)$ is the radiance predicted by the BRDF model utilizing parameters x and geometric vectors n, l, v . A total of 1152 measurement samples consisting of 48 view angles and 24 source directions per view are utilized. The views are clustered about the 30, 60, and 90 degree principal elevations where each camera image produces 16 views (a 4x4 grid) of the sample.

Functions that minimize RMS error give the best reconstruction for a scene, but not necessarily the best explanation of the physics. For example, a noisy reading or an impurity (geometric and material) might result in a specular spike in an otherwise diffuse material. A purely numerical fit may select a Phong model for a dark material on the basis of a large perceived brightness at 0 phase angle that poorly explains the presence of this lobe in other non-mirror views. Therefore, a secondary metric, Pearson correlation is also generated to validate the fit as it is invariant to scale and also measures the smoothness of the results:

$$P = \frac{\sum (E_i - \bar{E})(I_i - \bar{I})}{\sqrt{\sum (E_i - \bar{E})^2} \sqrt{\sum (I_i - \bar{I})^2}} \quad (5.16)$$

where \bar{E}, \bar{I} denote the mean of the observed and predicted radiance at each sample. Correlation values range from -1 to 1, where values near 1 indicate strong correlation. The selection of best BRDF explanation for the data is given using the following formula (5.17). The terms are scaled to give RMS error greater weight, but in general the two scores coincided strongly.

$$score = \varepsilon_{RMS} + 0.5 \times P \quad (5.17)$$

The RMS error and correlation coefficients used in scoring are detailed in fitting tables in the next sections. A third metric “% Error” (**max** error as a fraction of true value) is also included for reasons of human readability, though it does not always strictly coincide with RMS error.

The transformation from numerical albedo (which is specific to the experimental setup) to true albedo is performed by comparing results to the reference reflectivity of Spectralon, which is given as 0.99.

$$\rho_{true} = \rho_{raw} \times \left(\frac{0.99}{\rho_{spectralon-empirical}} \right) \quad (5.18)$$

Thus, multiplying the raw parameters in the next section by 3.3 (found empirically) gives scaled, real-world units. The next sections describe the BRDF fitting results of all the test materials and relevant background information.

Best-fits of each of the five BRDF models is shown visually by rendering on a test model of Eros, a near-earth asteroid. This model is physically relevant to the domain and possesses sufficient surface features to discriminate reflectance detail from a single view.

5.1.4.1 CMU-1 Simulant

CMU-1 is an optical lunar regolith simulant produced specially for this thesis. CMU1 was developed to be economical, expendable, and storable in laboratory environments lacking special ventilation. As such, it comprises only common inert and nonhazardous materials and can be easily mass produced (it is about 1/100 the cost of JSC-1A). The intent of these experiments is strictly optical (i.e. mapping, shape from intensity), so CMU-1 approximates regolith appearance. The simulant is not designed with regard to any mechanical properties. Terrains covered with CMU-1 are described in detail in the later application sections.



Figure 19. (Left) Design of CMU-1 lunar regolith simulant. The mixture ratio of coal and limestone dust is found by color and albedo matching against a known target under ambient illumination. CMU-1 (smaller-right) is qualitatively very similar to the fine portion of JSC-1A (larger-right), a NASA developed simulant.

CMU-1 is a mixture of coal and limestone in a 1.72:1 ratio. The constituents are pulverized such that 80% of the material by volume passes a 75 micron sieve. This is comparable to the “fine” portion of JSC-1A, a NASA developed simulant, in which 50% of the aggregate passes a 75 micron sieve. Both these materials are slightly hygroscopic, which contributes to low aerosol suspension and dispersion, despite fine granularity. The mixture ratio is the result of matching a 7% mean lunar (visible spectrum) Bond albedo [Willey 1976; Russell 1916] given these two reflectively dichotomous materials. Consideration of the spectral composition of CMU-1 is beyond the scope of this text; however the aggregate tristimulus color (as perceived with CIE 2° standard observer) is very nearly the same as the mare regions regolith, with a slight red tinge. Figure 19 illustrates the process of color and albedo matching to create CMU-1 and its qualitatively similar appearance to JSC-1A.

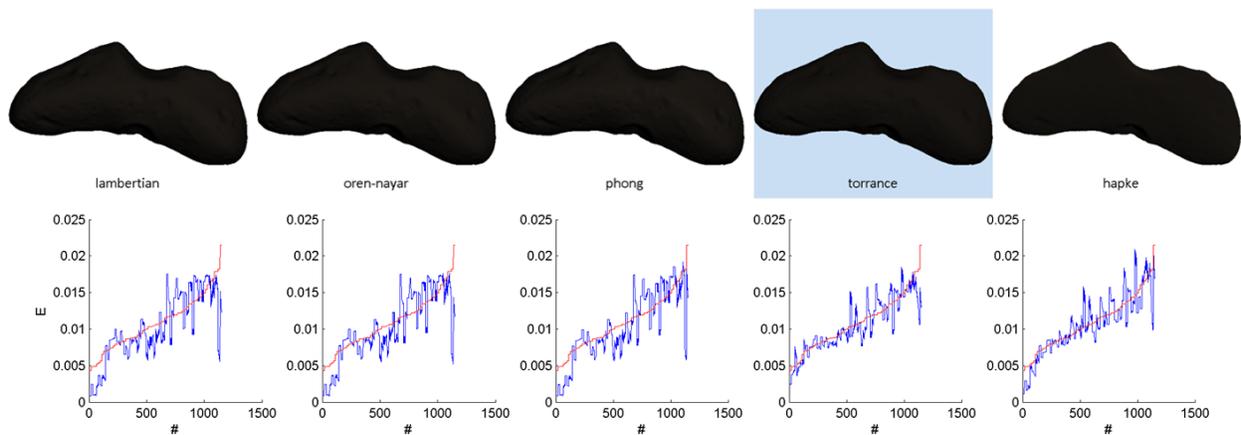


Figure 20. CMU-1 BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.

Indeed, gonireflectometry results show that CMU-1 and JSC-1A have strong quantitative similarities reflectively and perceptually. CMU-1 is best explained by the Torrance-Sparrow BRDF, which is the microfacet mirror model, with an error of 8.91% and a correlation of 0.864. The assumption of microfacets has a clear grounding in reality, as the sample consists of dust particles. There is moderate noise in the radiance curves, which is a consequence of the high-ISO camera setting required to image materials of low albedo. However, the recovered parameters are reasonable: the 6% empirical bulk albedo is very nearly the 7% of lunar regolith.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	13.3	0.295	0.796	$\rho_d = 0.022$
Oren-Nayar	13.4	0.295	0.791	$\rho_d = 0.023, \sigma = 0.12$
Phong	12.8	0.294	0.821	$\rho_d = 0.022, \rho_s = 0.0027, \alpha = 28.5$
Torrance	8.91	0.162	0.864	$\rho_d = 0.019, \rho_s = 0.012, r = 1.54, F_0 = 0.26$
Hapke	9.15	0.207	0.884	$w = 0.10, b = 9.3e-5, c = 0.93, B_0 = 0.23, h = 0.92$

The Torrance model also exhibits a statistically significant lower error than the other BRDFs. This error is about 50% lower than the Lambertian model, though only about 4.3% in terms of absolute error. Thus, CMU-1 is not particularly well represented as a diffuse material. However, in consideration of the intrinsic noise of the sensor and compared to Lambertian explanations of the other materials, it may be a sufficient approximation, particularly at non-glancing, non-mirror incident angles.

The Hapke BRDF is also a close fit to the data, albeit with several peculiarities. It features a slightly greater error (25% RMS, significantly less than the other BRDFs) than the Torrance model, but an even stronger correlation (2% greater). This is a promising development for CMU-1 simulant, as it is well known that Lunar regolith can be explained by the Hapke model. However, there are questions about the soundness of the recovered parameters, given that the Hapke BRDF is a strongly underconstrained model for fitting. For example, the phase asymmetry parameter b is very nearly zero. While this is technically possible in real materials, it is highly unlikely given prior evaluation of such lunar surface materials, which places a lower bound for this parameter around 0.05 [Aurelien Cord, 2003]. Likewise, the single scattering albedo w is about 60% lower than the typical value for dark lunar terrain at 1.6.

5.1.4.2 Coal Dust

Coal is a sedimentary rock formed by high temperature and compression of dead vegetation over hundreds of millions of years. Its primary constituent material is carbon and is perceptually a very dark substance. Coal is burned for heat energy and its use as a fossil fuel means mining is ubiquitous. Mines occur in seams, which are geological layers of near-homogeneous material in the ground. As such, the exposed coal represents a large portion of the interior surface in underground mines. The optical and geometrical regularity of these artificial voids is of particular interest in this work for vision application prompting characterization.

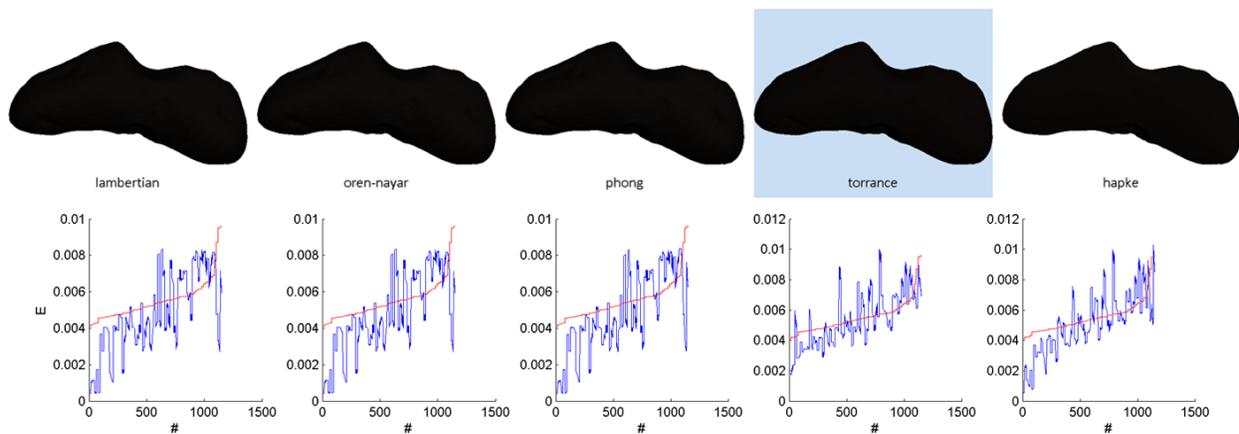


Figure 21. Coal Dust BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.

This work characterizes a sample of 75 μm , pulverized bituminous coal dust from the Pittsburgh seam. In addition to being easier to handle in this experimental setup, coal dust is pervasive in mines and a significant layer builds on walls as a result of active mining. Though pulverizing changes the surface reflectance properties of the material, it is also possible to infer the properties of the solid, particularly bituminous coal, as it is weakly coherent sediment.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	20.5	0.367	0.530	$\rho_d = 0.0104$
Oren-Nayar	20.5	0.367	0.530	$\rho_d = 0.0105, \sigma = 0.000277$
Phong	20.4	0.367	0.544	$\rho_d = 0.0104, \rho_s = 0.00037, \alpha = 23.4$
Torrance	13.7	0.248	0.665	$\rho_d = 0.0082, \rho_s = 0.0077, r = 1.43, F_0 = 0.33$
Hapke	16.6	0.307	0.661	$w = 0.054, b = 7.1e-5, c = 0.82, B_0 = 0.17, h = 0.16$

Results show that the appearance of coal dust is best explained by the Torrance model with statistical significance. The data is quite noisy, the consequence of an extremely dark material; thus, the error is fairly large and the correlation only moderate for all BRDFs. Though, the recovered parameters are sensible and consistent with classical sources. The albedo of 3% is about half the classical 6% value for a solid piece of coal. This is not unreasonable considering the rough, particulate nature of the sample. The recovered Fresnel coefficient (F_0) of 0.331 places the material squarely in the insulator category, between ruby and diamond; a determination that also appears sound.

Coal is not particularly well-explained by either the Lambertian, Oren-Nayar, or Phong reflectance models (7% greater error or 1.5x), which all parametrically reduced to Lambertian in optimization. Thus, unless an algorithm specifically handles Torrance-like specularity, a Lambertian assumption will suffice as well as the other three common models. The Hapke BRDF fares better (about 50% better fitting), but is still quite erroneous at glancing angles.

5.1.4.3 Concrete Gunite

Gunite (also known as shotcrete) is a form of spray concrete commonly used as a structural stabilizer and coating for surfaces in many terrestrial voids. Concrete for gunite varies, but usually comprise cement, aggregate and reinforcing fibers. The mechanism of spraying produces a rough surface finish that is distinct from typical concrete texture. The sample is a solid slab of Gunite which was cut from the Bruceton Research coal mine in Pittsburgh. Most of the mined-out surfaces of the Bruceton mine consist of this material.

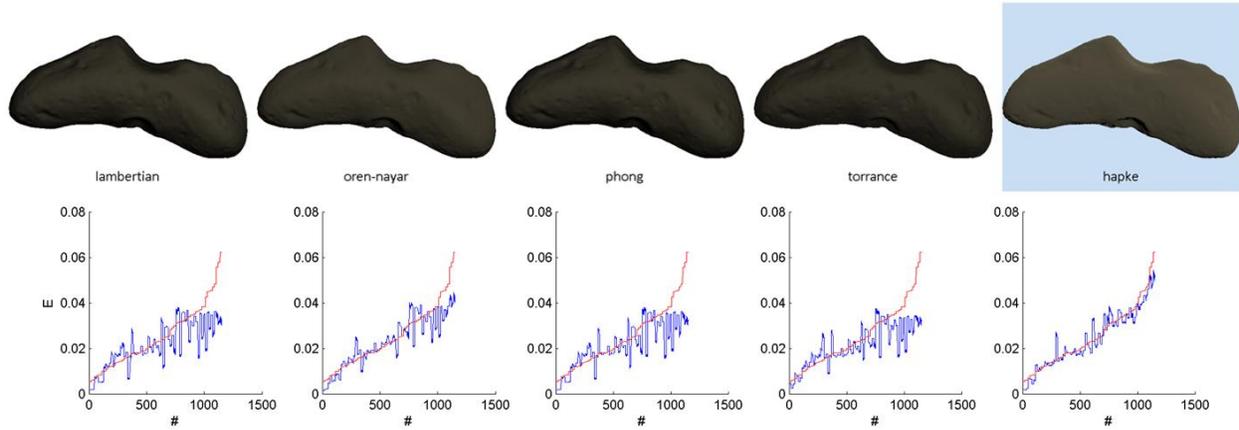


Figure 22. Gunite BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.

The constituents of Gunite are diffuse materials, and the amalgam is more so as a result of the macroscopically rough surface. As such, the Hapke BRDF is, quite reasonably, the only model with adequate complexity to fully describe the complex reflectance processes at play. The material is strongly backscattering ($c = 0.98$) and the surface can be seen from the point of the source as a series of concavities which mitigate scattering in the forward direction. Both particle phase parameters are consistent with a rough, high density material [Sato, et al 2012]. The total opposition effect is muted and broad as expected. The Oren-Nayar model also fits the data well, though the microfacet assumption most likely underrepresents the amplitude of surface texture. The high estimated roughness value from this model further corroborates the assumptions of surface geometry.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	14.3	0.302	0.781	$\rho_d = 0.046$
Oren-Nayar	10.1	0.241	0.891	$\rho_d = 0.075, \sigma = 0.85$
Phong	14.3	0.302	0.781	$\rho_d = 0.046, \rho_s = 2.5e-8, \alpha = 26.1$
Torrance	14.5	0.274	0.788	$\rho_d = 0.044, \rho_s = 0.012, r = 3.34, F_0 = 0.37$
Hapke	5.89	0.206	0.963	$w = 0.18, b = 0.28, c = 0.98, B_0 = 0.065, h = 0.15$

Overall, the error of the Hapke model is about 25% lower than Oren-Nayar and 50% lower than the others. The Lambertian, Phong and Torrance models are all basically equivalent in error and parameter values. A Lambertian BRDF does correlate moderately highly with the data, but the absolute sample error, which is modulated by the macroscopic displacement and self-shadowing of the surface, is quite

high. The corollary is that the Lambertian assumption clearly does not suffice for per-pixel vision approaches; however, it may be more appropriate in recovering trends aggregated over many surface cavities.

5.1.4.4 Granite

Granite is an abundant igneous rock that is widely distributed in the earth's crust. It comprises many minerals such as quartz and feldspar that give a speckled look. Component and hence bulk colors vary broadly. Many natural caves and voids are granite, particularly where water has eroded softer rock between large slabs. The Granite sample analyzed is a salt and pepper type, but is perceptually gray. It is slightly weathered, with a coarse, but uniform surface.

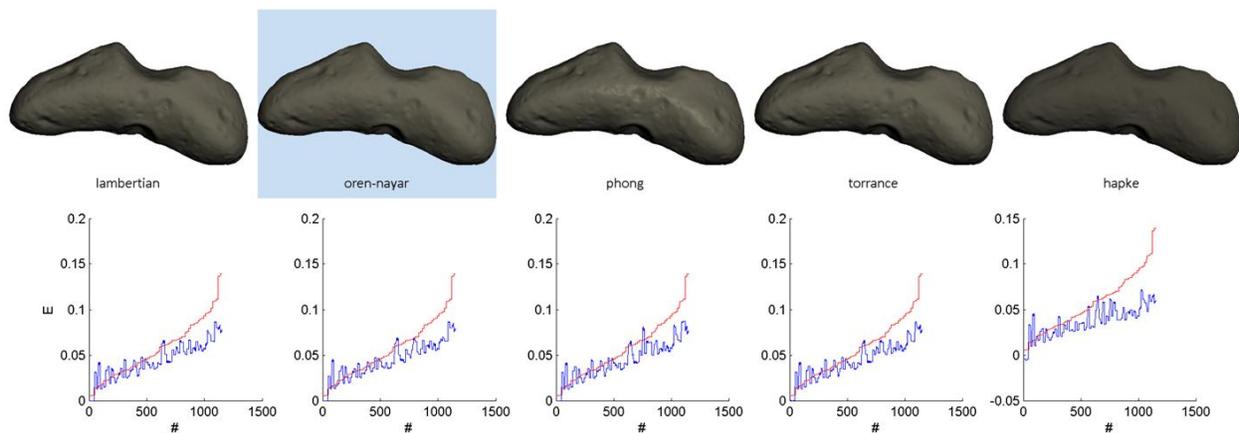


Figure 23. Granite BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.

The Granite sample was found to be Oren-Nayar in appearance, though it is only insignificantly non-Lambertian along with the other reducible models. Noise levels in the data were low; however, there is a bias trend in fitting against any of the BRDFs that contributes to a medium-high absolute error. This error is believed to be the result of a slight, but nontrivial, convexity of the sample. It is likely that given a more planar sample, the material can be classified as very strongly Lambertian, as correlation is very high. The Hapke model does not fit the data, indicating that there is little subsurface scattering.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	14.3	0.400	0.903	$\rho_d = 0.084$
Oren-Nayar	14.1	0.395	0.908	$\rho_d = 0.091, \sigma = 0.24$
Phong	14.3	0.398	0.889	$\rho_d = 0.083, \rho_s = 0.0191, \alpha = 18.3$
Torrance	14.3	0.400	0.903	$\rho_d = 0.084, \rho_s = 8.2e-8, r = 0.94, F_0 = 0.56$
Hapke	20.1	0.572	0.816	$w = 0.61, b = 0.99, c = 0.078, B_0 = 0.0074, h = 0.91$

5.1.4.5 Gray Sandstone

Sandstone is a soft sedimentary rock formed from silica (sand), calcium carbonate and other minerals. The appearance of sandstone varies widely and it may take on a gray, tan, pink and yellow color. Even combinations of these colors in the same sample are possible as the material typically exhibits an anisotropic banding as a result of sedimentation.

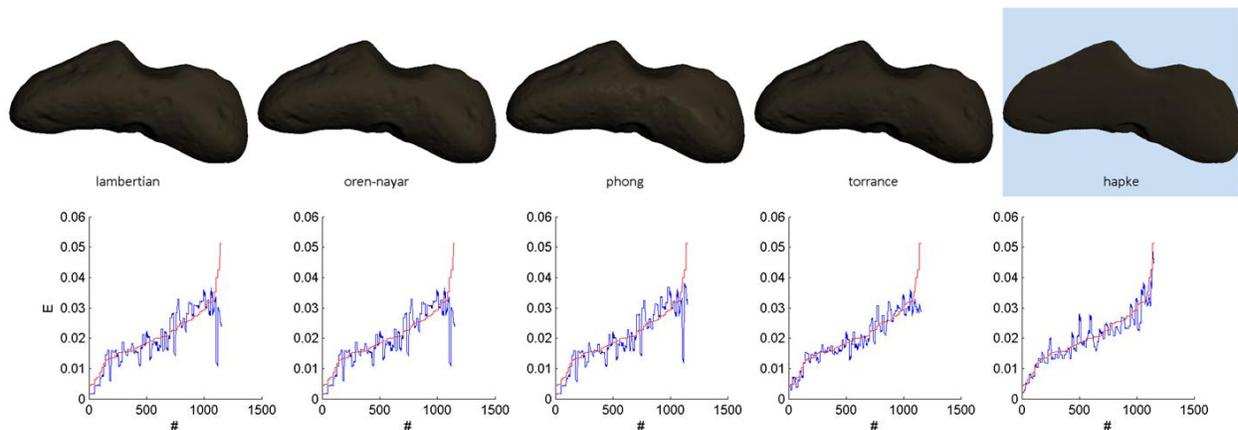


Figure 24. Sandstone BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.

The sample used for testing is a slab of gray sandstone collected at Walker's Mill cave in Pennsylvania. The walls of this cave consist entirely of sandstone slabs which are homogeneous in appearance. The surface of the sample is coarse but uniform, and speckled with minute reflective crystals. Intrinsic anisotropic banding is perceptually undetectable; thus, it is appropriate to utilize the 3-DOF gonioreflectometer for complete characterization.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	11.5	0.248	0.784	$\rho_d = 0.041$
Oren-Nayar	11.5	0.248	0.784	$\rho_d = 0.041, \sigma = 1.8e-5$
Phong	10.2	0.244	0.837	$\rho_d = 0.041, \rho_s = 0.0090, \alpha = 25.4$
Torrance	7.82	0.160	0.893	$\rho_d = 0.036, \rho_s = 0.010, r = 0.90, F_0 = 0.35$
Hapke	6.11	0.161	0.939	$w = 0.21, b = 0.26, c = 0.23, B_0 = 0.13, h = 0.94$

Sandstone is optimally described by the Hapke model with very high correlation and low error which is near the threshold of sensor noise. In terms of physical properties (b, c, h) explained by the model, it is a densely compacted (high h), agglomerate of medium surface roughness (medium-low b and c). Of the common BRDFs, it is also well explained by the Torrance model which has statistically insignificant RMS error and about 5% lower correlation in comparison. The surface is distinctively microfaceted and specular (this can be seen in the crystalline nature of the speckling).

The metrics for the remaining BRDFs are all numerically similar: 50% greater RMS error than the Hapke model, though moderately low in absolute terms, and medium-high correlation. Like Gunite, sandstone is clearly not Lambertian, though the absolute error is low enough that it may be acceptable to make the assumption in many cases.

5.1.4.6 JSC-1A Simulant

JSC-1A is lunar regolith simulant for titanium-deficient mare (dark flat plains) regions. It is chemically similar to Apollo returned lunar regolith samples; the primary constituent is a basaltic ash with high glass content mined at the San Francisco volcano fields in Arizona [Ray, et al 2010]. Particle size distribution is also designed to mirror that of Apollo soil samples. The raw material is coarsely sieved and then comminuted in an impact mill to generate the appropriate sizes. As such, constituent particles are particularly jagged reflecting the meteoritic weathering of the moon. Grains range from 1 μ m to a maximum of 1mm and an average size of approximately 81 μ m [McKay, et al 1994].

As a NASA developed simulant, JSC-1A (unlike many of the other materials documented here) has been comprehensively characterized; mechanically, chemically and photometrically. Spectroscopic analysis in particular, has been done with high rigor, utilizing dedicated instrumentation in highly controlled environments. Reflectance curves have been recovered across the spectrum of visible and invisible wavelengths. This work does not seek to reproduce these extensive results. Rather, a holistic,

perception centric approach is taken to augment prior work, which is deficient in this area. For example, BRDF results are either presented as wavelength-indexed lookup tables [Johnson, et al. 2008] or collections of optimized Hapke parameters [Helfenstein, et al. 1987; Cord, et al. 2003]. These formulations do not lend themselves to modern computer vision which requires linearly separable BRDFs, assumes albedo is a constant intrinsic property, and operates in a contrived RGB space.

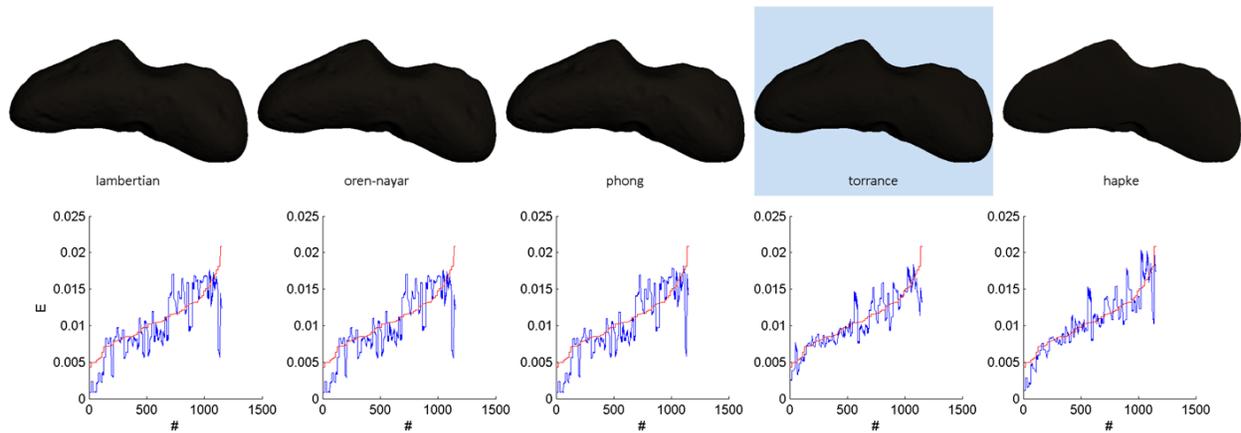


Figure 25. JSC-1A BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample projection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.

JSC-1A was produced to enable comprehensive and destructive scientific testing not possible with genuine lunar samples. However, JSC-1A is expensive, limited in availability, and no longer produced. The sample used for testing was loaned from Glenn Research Center. Though it was kept in a sealed container, prior use and replacement measures did not specifically mitigate moisture and other possible airborne contaminants.

Experimentation shows that JSC-1A is similarly representable as both a Torrance and a Hapke material. Torrance fitting is lower in absolute error, while the Hapke model presents higher correlation with observed values. Overall noise is moderately low for the raw data. As a weakly cohering particulate, JSC-1A reasonably satisfies both the microfacet Torrance and the multiple scattering Hapke assumptions.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	13.7	0.298	0.787	$\rho_d = 0.021$
Oren-Nayar	13.7	0.298	0.790	$\rho_d = 0.022, \sigma = 0.062$
Phong	13.3	0.297	0.810	$\rho_d = 0.021, \rho_s = 0.0022, \alpha = 27.9$
Torrance	8.63	0.159	0.872	$\rho_d = 0.018, \rho_s = 0.010, r = 1.49, F_0 = 0.33$
Hapke	9.17	0.208	0.890	$w = 0.10, b = 2.6e-5, c = 0.87, B_0 = 0.17, h = 0.98$

However, as with the CMU-1 material, it is unlikely that recovered parameters in the Hapke model are physically valid. The particle phase parameter (b) is nearly singular here; while prior work argues c should rarely exceed 0.5 for lunar type terrains [Helfenstein, et al. 1987]. The value of the single scattering albedo (w) is plausible, but the distribution of opposition effect (h) is much too broad and amplitude (B_0) too muted. Thus, it is possible that the single and multiple scattering processes tied to w are sound, but there is insufficient angular sampling to estimate the opposition effect with any numerical stability. The Torrance model is a much more plausible explanation of the appearance, with physically reducible parameters within acceptable ranges.

The other BRDFs are poor explanations for JSC-1A, with almost 100% greater error and 10% lower correlation. However, this difference is reduced somewhat by considering the low total error, and increased measurement noise caused by the material's low absolute reflectivity.

5.1.4.7 Limestone Dust

Limestone is a light gray sedimentary rock consisting of calcium carbonates such as calcite and aragonite. It is formed primarily by the accumulation of marine skeletal fragments (bones and shells) on the sea floor. Limestone comprises 10% of all sedimentary rock on earth and is optically important in terrestrial underground environments due to its abundance. Most caves are limestone as it is easily eroded by hydraulic processes, due to its soluble nature. It's also a common surface material in artificial voids where it is mined, and also due to its nonreactive properties.

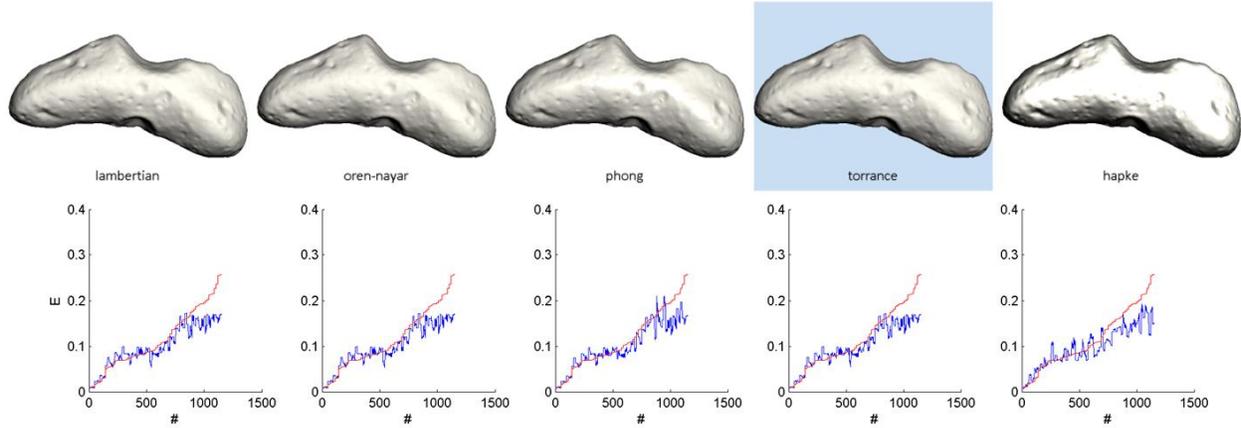


Figure 26. Limestone BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.

Limestone serves as a building block of many other materials, some which are explored in this work. Industrial uses are pervasive: in powdered form, agricultural lime is used to reduce the pH of soils, in the creation of concrete, as an explosion suppressant in coal mines, and even as an edible source of calcium. As a crushed aggregate, it is used structurally and as gravel. Lastly, blocks of limestone are utilized in building construction. A sample of pulverized limestone dust is characterized in this work.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	11.4	0.201	0.945	$\rho_d = 0.22$
Oren-Nayar	11.4	0.201	0.945	$\rho_d = 0.22, \sigma = 0.0010$
Phong	11.1	0.191	0.934	$\rho_d = 0.21, \rho_s = 0.047, \alpha = 12.9$
Torrance	11.5	0.201	0.945	$\rho_d = 0.22, \rho_s = 0.002, r = 1.5, F_0 = 0.24$
Hapke	12.9	0.252	0.928	$w = 0.65, b = 0.37, c = 0.50, B_0 = 0.22, h = 1.2e-6$

Results show that limestone is almost certainly a Lambertian material. Though the Torrance model scored marginally highest, it and the other reducible BRDFs are statistically inseparable from the Lambertian model with more than 99% of the energy being diffuse. Correlation is high and error is low in this determination. The Hapke model is comparatively neither a fit to the data or the parameters as the compaction parameter h is numerically zero.

5.1.4.8 Ochre Paint

This sample is a yellow-orange colored matte paint used on the Gretag-Macbeth Colorchecker™ SG (semi-gloss), a color calibration target. This material is not found in the environments of interest, but serves as a contrast to the other materials characterized. Most of the planetary materials are matte shades of dark gray and there is great difficulty distinguishing them from known quantities. The unique off-white color and visible specular reflection demonstrate that the experimental setup is capable of detecting a variety of intrinsic BRDFs with fidelity.

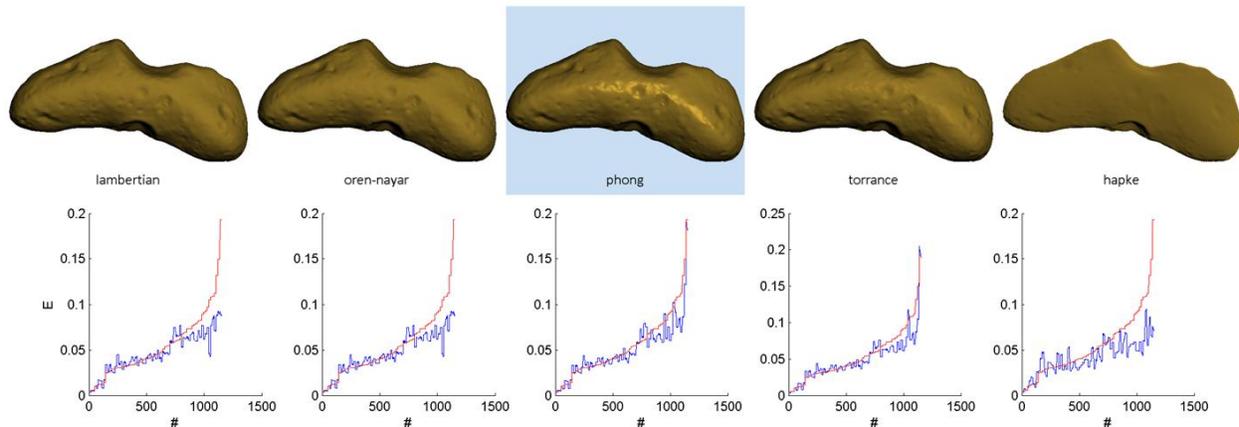


Figure 27. Ochre Paint BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue.

The ochre semi-gloss paint is a moderately specular material and is described equally well by both the Phong and Torrance models, although the Phong BRDF has an insignificantly higher correlation with observed data. The Lambertian model is capable of describing the paint at non mirror angles (it is a linearly separable BRDF), with low error and high correlation. Given the relative sharpness of the specular lobe ($\alpha = 18.6$), which is equivalent to a small highlight, a Lambertian assumption would be valid in a simple illumination environment with overwhelming probability.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	10.3	0.216	0.889	$\rho_d = 0.096$
Oren-Nayar	10.3	0.216	0.889	$\rho_d = 0.096, \sigma = 0.0006$
Phong	6.80	0.192	0.942	$\rho_d = 0.093, \rho_s = 0.049, \alpha = 18.6$
Torrance	6.98	0.192	0.940	$\rho_d = 0.095, \rho_s = 0.018, r = 0.13, F_0 = 0.041$
Hapke	13.5	0.330	0.804	$w = 0.31, b = 0.17, c = 0.94, B_0 = 0.003, h = 0.95$

5.1.4.9 Spectralon

Spectralon™ is utilized for radiometric calibration of the experimental setup; it is the most diffuse material known. Its diffusive properties are the result of a fluoropolymer structure, which produces isotropic multiple subsurface reflection. The material is spectrally white and exhibits over 99% reflectivity, which makes it ideal as a calibration sample. The known diffusivity and reflectivity are exploited to recover two parameters. These parameters are (1) a scalar conversion factor from perceived pixel radiance to unit reflectivity values and (2) a calibration of the irradiance incident on the sample from each light source as a function of distance. The second measurement also compensates for manufacturing uncertainty in the radiance and errors due to the small areal nature of each source. The procedures for these calibrations are discussed in Section 5.1.2.3 (Calibration).

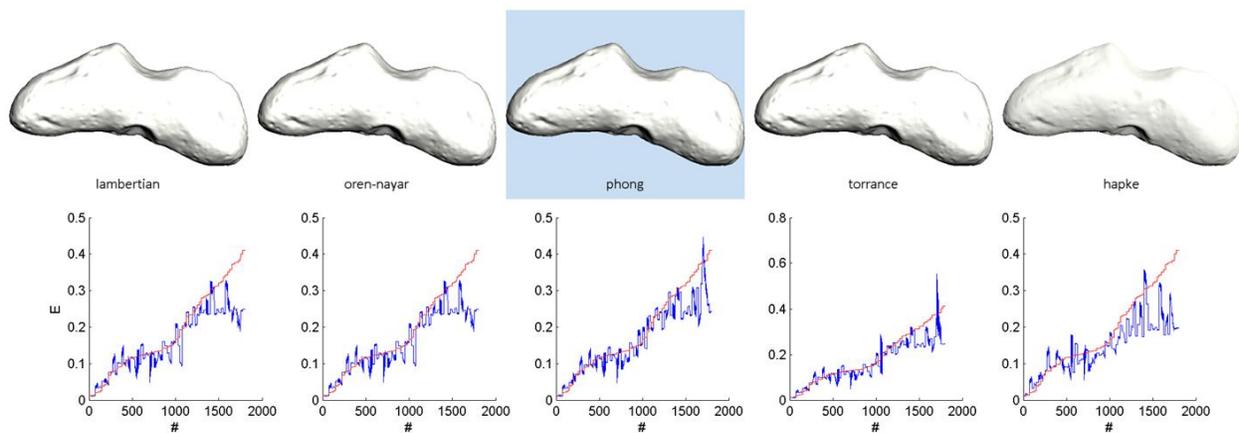


Figure 28. Spectralon BRDF Fitting Comparison. Reflectance functions rendered on a test model (top) and Per-sample reprojection errors for each BRDF (bottom). Observed radiances in red are sorted by magnitude; predicted data is in blue. There is bias error for high-intensity measurements due to semi-saturation. This and other errors prompt regularization of calibration parameters, regardless of “known” sample values.

Data from this reference sample is included here to illustrate typical measurement data and error sources on a known material. The sample is clearly Lambertian as expected, with the insignificantly

better Phong fit within 1% difference. This corresponds to expectation, but is likely the result of overfitting to noise. While correlation is necessarily high and overall noise minimal, sources of error are present and detectable in this characterization. These errors assumed for other samples and mitigated in a variety of ways. Discussion of this process occurs in the next section.

BRDF	% Error	RMS Error	Correlation	Raw Parameter Values
Lambertian	13.7	0.245	0.920	$\rho_d = 0.31$
Oren-Nayar	13.7	0.245	0.920	$\rho_d = 0.31, \sigma = 0.0002$
Phong	11.4	0.241	0.937	$\rho_d = 0.30, \rho_s = 0.024, \alpha = 6.33$
Torrance	12.0	0.239	0.932	$\rho_d = 0.31, \rho_s = 0.051, r = 0.19, F_0 = 0.14$
Hapke	18.2	0.349	0.865	$w = 0.77, b = 1.4e-7, c = 0.83, B_0 = 8.3e-11, h = 0.70$

5.1.4.10 Error and Accuracy

The Spectralon data illuminates two possible physical sources of error. The first is saturated measurement of very bright or dark objects. Camera radiometric curves are not accurate near the pixel limits (0 and 255). Specifically, when RGB channels are combined to a single irradiance value, this may have the effect of magnifying noise instead of a reduction as intended. Utilizing a dead-band both reduces dynamic range and does not significantly address the problem of edge values. In classic HDR imaging, a probabilistic approach is taken where Gaussian weights are used to emphasize “well-exposed” values in combining an exposure bracket. However, for particularly bright objects, near-saturated measurements still dominate given a fixed number of exposures to consider. A well-exposed image may not exist for any given set of images in a bracket. It quickly becomes infeasible to take many more images – for example, in an exposure adjustment loop – for the entire range of intensities produced by the incident angles. Fortunately, Spectralon and coal were the only materials where saturation was a significant concern.

The second source of error is due to non-planarity of the sample leading to misassumption of the surface normal and extreme modulation of perceived intensity from self-shadowing. The Spectralon used is macroscopically flat but worn; there are visible indentations on the surface (the material is quite soft). These indentations, as well as convex features on other samples, are shadowed by glancing sources. These glancing sources are also physically nearest to the sample, creating a complex relationship. While BRDFs are capable of addressing self-shadowing in a microfacet sense, these

assumptions break down with macroscopic deviations; those that approach the size of 1 measurement sample. These surface errors could explain the BRDF preference for a specular spike in fitting the Spectralon data.

Manual inspection provides the best tool for dealing with data acquisition errors. Spectralon, a very bright object, was imaged with decreased exposure (increased shutter speed), while coal was imaged with increased ISO sensitivity (there is a maximum integration time allowable due to physical heating of the sources). However, raising the ISO, introduces additional measurement noise.

The approach to mitigating systemic errors involves certainty of large numbers. Regularization of the distance polynomial in irradiance compensation prevents overfitting of saturated regions of the Spectralon sample in calibration. This is also the primary reason for noise in the well-exposed regions of the radiance curves despite Spectralon being an “ideal” material. Minimal self-shadowing is simply tolerated. Given the 1152 measurements amortized over 16 measurements by 72 images, any individual errors would have negligible effect on the total data. Particularly egregious cases, however, are simply removed from consideration.

Analysis of the Spectralon data shows the **overall intrinsic error of the gonireflectometer to be approximately 5%**, a tolerable value for the purposes of this work. This value does not include additional sample-specific errors that may result. These may include the cameras’ ISO-dependent, integration noise or physical errors such as inaccurate sample placement or airborne dust in the measurement volume. These other error sources are also manually detected and removed if possible, but it is necessary to accept some noise as unmitigable.

5.1.5 Discussion

A summary of perceptual measurements are given in Table 5. The planetary materials characterized ranged in diffuse albedo from a low of 3% for coal dust to a high of 68% for limestone dust (not including the reference materials). Typical terrestrial underground albedos are in the 10-20% range, while planetary regolith was much darker at 6%. Illumination-carrying robots are thus at a significant sensing disadvantage in planetary environments, given the same output power. With open, planar geometry precluding illumination by interreflection, it is prudent to leverage natural illumination for imaging - particularly sources that may be approximated by simple points (such as lunar sunlight).

Materials are overwhelmingly shades of gray, with slight red tinges in the regolith materials, coal and sandstone. This has little significance in reflectivity analysis, but may be useful for discriminating between soils and their bedrock.

Table 5. Table of Observed Color and Albedo Values

Material	sR	G	B	Albedo
CMU-1	0.65	0.58	0.49	0.06
JSC-1A	0.66	0.58	0.48	0.06
Coal Dust	0.69	0.54	0.46	0.03
Granite	0.61	0.60	0.51	0.25
Gunitite	0.64	0.60	0.48	0.14
Limestone Dust	0.59	0.58	0.56	0.68
Sandstone	0.67	0.59	0.45	0.13
Ochre Paint	0.77	0.61	0.19	0.28
Spectralon	0.58	0.58	0.57	0.99

* Materials are sorted by planetary, underground and reference types. Colors are given as normalized values on the unit sphere.

Reflectively, three of the materials characterized were found to be *definitively* Lambertian: limestone, granite and the reference spectralon. Materials with less than 1% difference in score between the best fitting BRDF and the Lambertian explanation are given this designation. The next class of *strongly* Lambertian materials features less than a 5% difference; this category consists of only the reference ochre paint. The class of *moderately* Lambertian materials - those with less than a 50% score difference - include coal, concrete, and sandstone. Lastly, the class of *slightly* Lambertian materials have as much energy in the Lambertian component as not (<100% difference). This category includes the regolith simulants CMU-1 and JSC-1A. There are no materials in this study that can be classified as definitively non-Lambertian (>100%).

In this classification, relative score differences are used for comparison. When the absolute error and intrinsic noise of the data is considered, **most materials can be tolerably represented as Lambertian** with less than 15% error and greater than 0.8 correlation. There are several important corollaries from this determination. Firstly, variegation (spottedness) is sufficiently approximable with a single physical albedo value for each material. Secondly, aggregate materials, regardless of composition can be represented by a single bulk BRDF with high accuracy, i.e. all fits have tolerable absolute error. In multi-term BRDFs this means that the diffuse component is either Lambertian or the material is strictly Oren-

Nayar. Oren-Nayar/Torrance hybrids, for example, are not significantly more expressive in this domain. Lastly, materials with purely specular BRDFs (zero diffuse albedo) do not exist.

Of the materials that are only moderately Lambertian or less, three of them are particulates best explained by the Torrance model, while gunite and sandstone are Hapke materials. While the Hapke BRDF is physical model of dusty surfaces, it proves difficult to invert and fit for these experiments. The three Torrance dusts, coal, JSC-1A and CMU-1, are numerically a good fit for Hapke, but parametrically nonsense. In fact, the Torrance model provides a better fit given the data, and is furthermore receptive to vision techniques. In materials that are indisputably Hapke (low error, parametrically reasonable), gunite is also well-explained by Oren-Nayar and sandstone by Torrance-Sparrow. Thus, it can be argued that explicit consideration of the Hapke model, while providing a low-error fit for many materials, is unnecessary in fully describing the gamut of planetary materials.

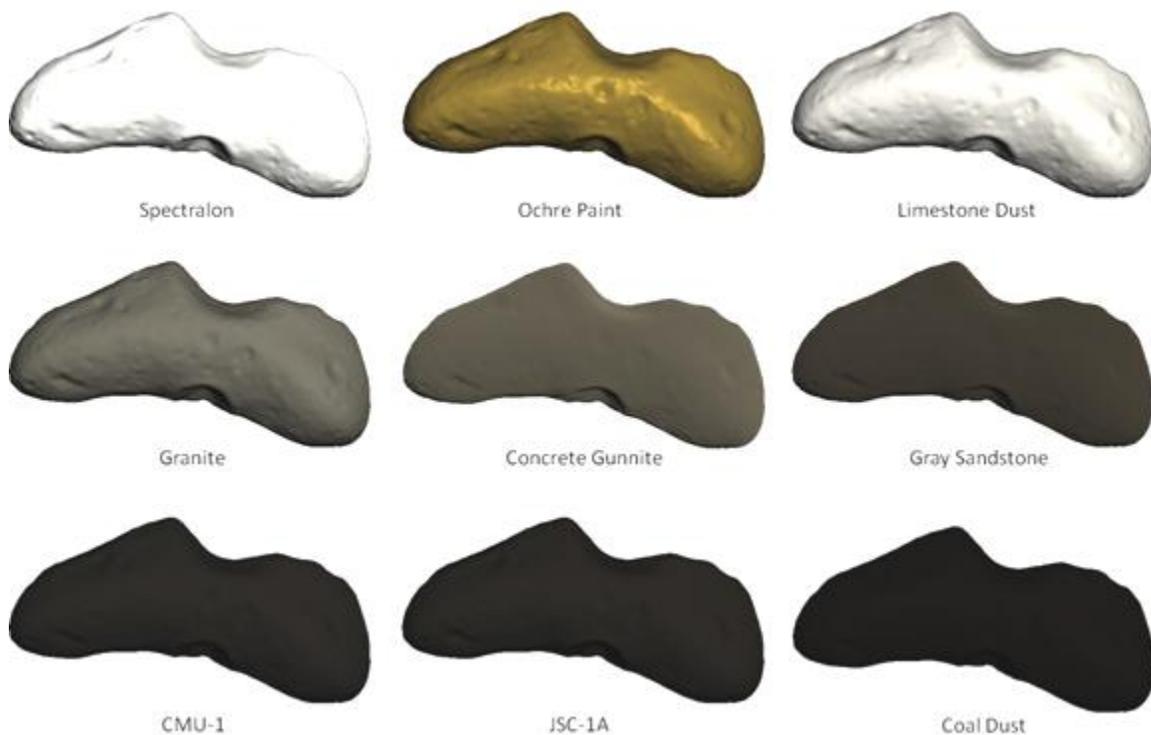


Figure 29. Rendering of Fitted Planetary and Reference Material BRDFs. Color, relative albedo, and reflectance functions are derived from experimental data; brightness is enhanced for print and screen viewing.

Considering only the four “traditional” graphics models: **limestone, granite and Spectralon are Lambertian; the semi-gloss paint is Phong; coal, sandstone, JSC-1A, and CMU-1 are Torrance and**

gunite is an Oren-Nayar material. As the paint is not a planetary material, just three BRDFs – Lambertian, Oren-Nayar, and Torrance – are sufficient to span the reflective space of the domain.

Rendering of these material BRDFs on the Eros asteroid model are shown in Figure 29. Rendering uses a single best-fit BRDF with optimized parameters, relative albedo and color estimates. Any detectable surface anisotropy, texture or variegation (i.e. banding of sandstone) are not utilized. Illumination is a single infinitely-far point source that is slightly tilted to the positive Cartesian quadrant from the camera center. Brightness has been enhanced for display of low-reflectivity materials. Polar slice visualizations of the BRDFs for each material are illustrated in the Appendix: 7.4.1.

5.2 Experimental Environments and Distributions of Materials



Figure 30. Typical Imagery from environments considered in characterization: (left) Bruceton Coal Mine, (center) macroscopic Lunar terrain and (right) Walker's Mill sandstone cave.

Surface appearance not only depends on the optical properties of materials, but also the frequency and macroscopic distributions of those materials. These distributions are functions of individual environments. Thus it is not sufficient to simply characterize common materials in order to understand appearance. The environments where they occur must be characterized for the manifestations of these materials.

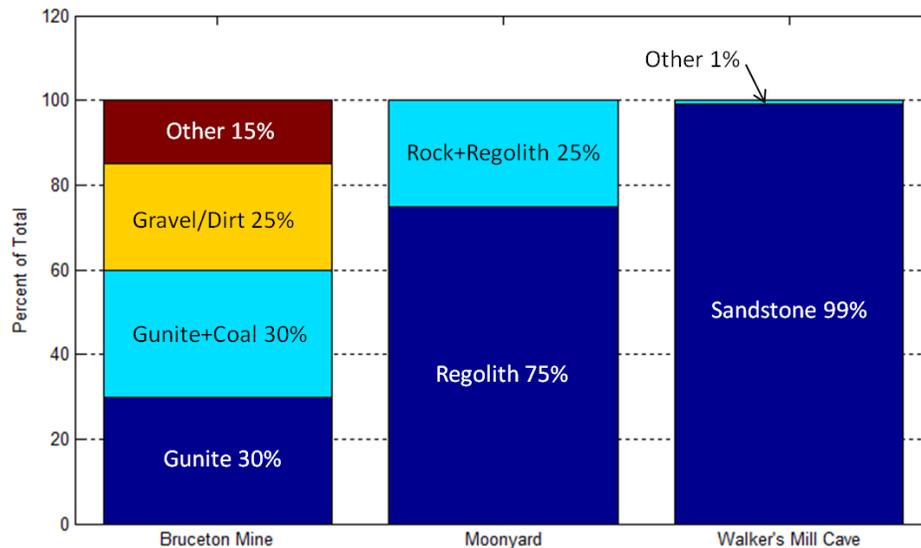


Figure 31. Approximate Prevalence of Surface Materials in Experimental Environments from qualitative analysis. Notes: the other category of Bruceton mine denotes artificial ceiling materials and mining equipment. Gunite+coal denotes a discrete variegated surface of these two materials. The rock+regolith category under the moonyard column denotes a variegated surface of regolith and rock.

This section considers aspects of the *barren* and *rocky* qualifiers which are functions of the macroscopic distribution. It is beyond the scope of this work to characterize a spanning set of planetary environments, thus only analysis of several representative environments used for experimentation is presented here. These environments are: an underground coal mine, a lunar analog terrain, and a natural cave. The nature of these environments and their optical similarity to the general planetary population is discussed below. Figure 31 summarizes the approximate distribution of materials as fractions of the total internal surface area.

Bruceton Coal Mine

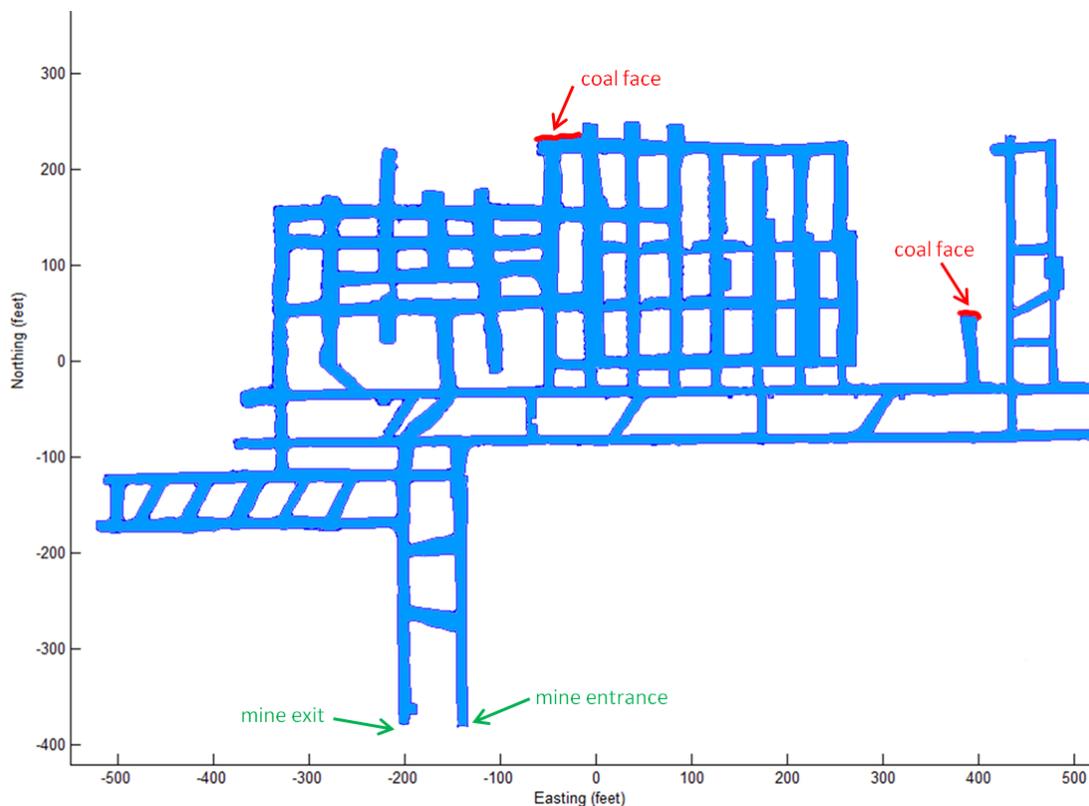


Figure 32. Overhead Map of Bruceton Mine, the primary experimentation environment for mobile robots in this thesis. This mine consists of two appearance environments, (1) Corridors consisting of Gunnite covered walls and ceiling and dirt floors which constitute over 98% of the surface area of the mine and (2) open coal faces which are less than 1%. Internal objects, such as mining vehicles constitute the remainder of surface area. Blue areas represent the open void of the mine within the solid strata.

Bruceton is a coal mine in Pittsburgh, located in a bituminous coal seam. It is the main experimental environment in this thesis, due to its accessible conditions and location. Bruceton is representative of

coal mines in many ways. Perhaps most significantly, it shares the similar macroscopic geometry of the room and pillar style (Figure 32). It is possible to describe the mine as consisting entirely of homogeneous corridors which meet in right-angle three or four-way intersections or dead ends. It also features compacted dirt floors, rail tracks in primary corridors, jarring plastic mine curtains and the occasional mine vehicle typical of most underground mines, not just the coal variety.

However, Bruceton is also a “research” mine, meaning that it is structurally well maintained in comparison to commercial coal mines and also that there is no active mining of coal. The appearance of Bruceton differs from commercial mines due to these circumstances. First, the roof is stabilized using several methods, including girders for structural support, gunite (a diffuse and uniform surfacing material) for stabilization and plastic mesh to catch spalled bits of materials. These measures are not each in effect over the entirety of the mine; however, there are no places where none of the measures are utilized. These roofing materials, particularly the plastic, are among the worst offenders of the diffuse and barren assumptions in the mine. In active mines, gunite and mesh are not as commonly used for economic reasons.

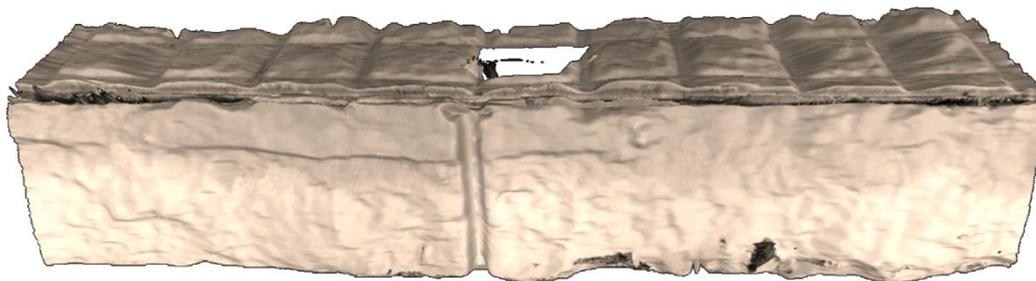


Figure 33. 3D mesh model of corridor in Bruceton Mine, generated by mobile robot.

The walls of the mine are almost entirely covered in concrete gunite which acts as a surface stabilizer and flame retardant for exposed coal. This condition strongly satisfies the vision assumptions discussed. While, active mines also spray the walls with limestone dust for the same purposes (and is likewise diffuse), gunite represents a different material with an extra level of uniformity due to its more permanent adherence to surfaces. The main corridors in Bruceton are extremely well maintained; however this decreases in areas with less frequent use. Noticeable surface variegation from cracked and broken coatings of old gunite, which exposes the underlying coal, is common in the many minor corridors.

The lack of active mining means that most of the exposed coal faces have been covered. Therefore, while Bruceston comprises two distinct environments like most mines - mined-out gunnite corridors and open coal faces – the latter represents only a small portion of surfaces. In fractional terms, this is not unlike active mines; however, as Bruceston is a small mine, coal faces typically represent a much greater explorable region. Active mining also creates pervasive coal dust that accumulates on surfaces driven by air currents. In mines with limestone dusting, the surface coating may be similar in appearance to CMU-1. There is practically no coal dust on surfaces at Bruceston.

CMU Moonyard

The CMU moonyard is a lunar analog terrain constructed in a 2m x 1.3m box. This terrain represents macro-scale features at a critical sensing distance for surface robots. A rock size distribution that is the average of all the lunar Surveyor sites (Figure 35) was randomly generated on terrain using procedural simulation. This was then used as a blueprint for realizing the terrain using analog materials, where fidelity of appearance was the major objective. A bed of limestone dust about two inches thick was used to create smooth undulations (these were not simulated). Crushed limestone, sieved through critical diameters of [2,4,8 ...] mm, were then hand-placed in the generated locations for all rocks larger than 2mm in diameter. The largest rocks were sized manually, and as diameter is an idealization, some oblong rocks were allowed to deviate significantly in the minor axis.

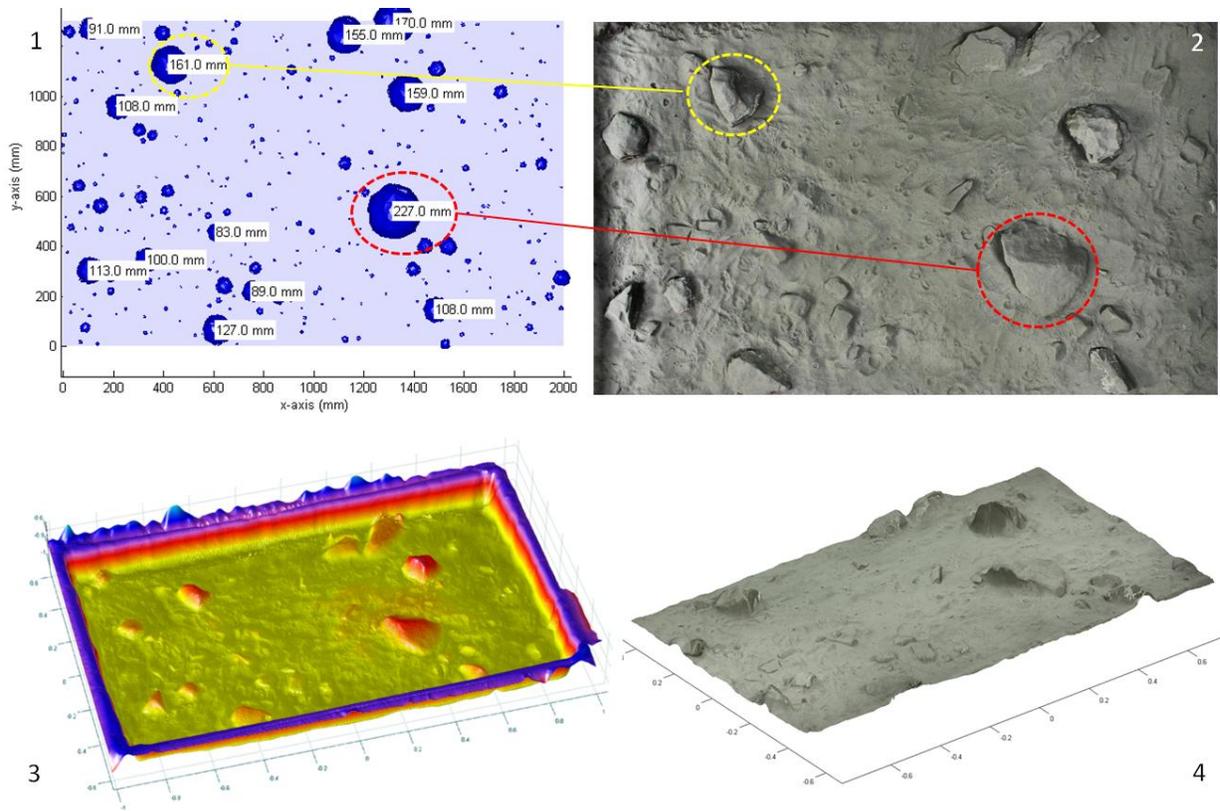


Figure 34. A simulated lunar rock distribution utilizing Surveyor data (top left) and realization in the construction of an artificial moon scene (top right). The bottom row illustrates a z-colored mesh model of the scene generated with LIDAR scanning (bottom left) and color-mapping of the terrain from DSLR images (bottom right).

A layer about 1mm thick of CMU-1 was then dusted across the entire surface, giving a mostly uniform covering²⁰. Micro craters were created by ballistic placement of pebbles less than 1mm diameter and in the dusting process. Figure 34 shows the ideal distribution of features generated in simulation and the actual placement of these features as-built.

²⁰ Some highly sloped edges reveal underlying rock.

Landing site	$K, \text{mm}^{-\gamma}/100 \text{m}^2^a$	γ^a	Number of particles measured
Surveyor I	5.0×10^5	-2.1 ₁	2192
Surveyor III	3.3×10^6	-2.5 ₆	1068
Surveyor V	$1.2_5 \times 10^6$	-2.6 ₅	2803
Surveyor VI	$1.9_1 \times 10^6$	-2.5 ₁	1766
Surveyor VII	7.9×10^5	-1.8 ₂	2077
Total			9906

^a $N = KD^\gamma, 1 \text{ mm} \leq D \leq K^{-1/\gamma}$, where N is the cumulative number of particles with diameter equal to or larger than D per 100 m^2 , and D is the diameter of particles in millimeters.



Figure 35. Distribution of Rock Sizes (left) at the Lunar Surveyor landing sites, from NASA Surveyor Project Final Report 1968. Creation of the moonyard utilized crushed and sieved limestone rocks of different diameter (right).

The moonyard enables convenient physical testing of lunar algorithms. While it is acknowledged that a single terrain and even the surveyor distributions themselves cannot account for the variance of lunar appearance, this is a sufficiently representative environment for testing. Space weathering through prolonged meteorite bombardment has covered the moon in a layer of regolith of appreciable thickness. This reduces total variegation, even as rocks of many materials appear on the surface and regolith composition is specific to regions of the moon. The mare regolith is closely matched in albedo and reflectance to CMU-1. Moreover, the frequency and distribution of rocks compares favorably to Apollo images. There is little in the way of *distinctive* features that may contribute to systemic overfitting to the moonyard where applicable; no individual features are considered. Lastly, training of lunar algorithms occurs mostly on simulated data in this thesis, which decouples results from physical demonstration²¹.

²¹ The moonyard is used to make generalizations about appearance, but these are not directly utilized in computation of approaches.

Walker's Mill Cave

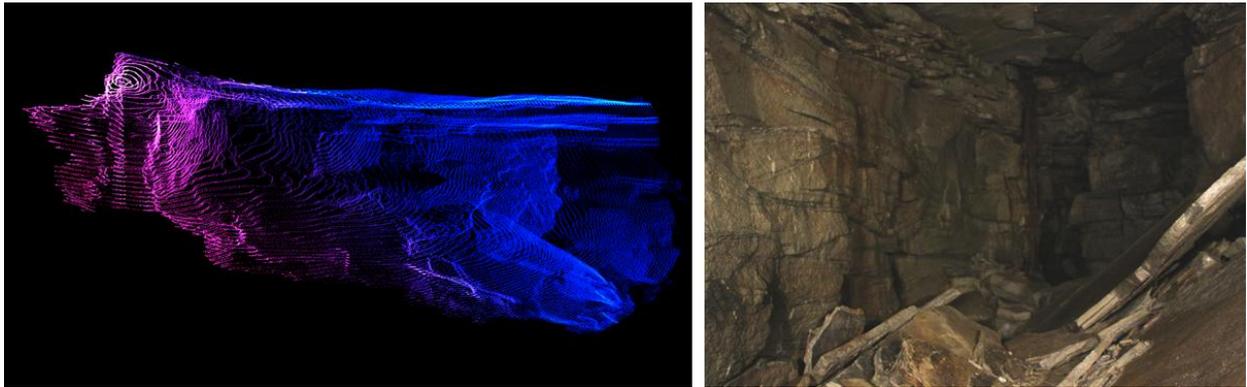


Figure 36. A partial LIDAR scan (left) and photo (right) of Walker's Mill cave showing the irregular geometry of fractured sandstone.

Walker's Mill cave is a natural sandstone cave in Pittsburgh. It is diminutive, with the main room no bigger than 2m tall, and about 11m long; various branches are too small to accommodate humans. The entrance to the cave is about 0.5m square and requires an almost vertical entry, making mapping of this cave a perfect application for robots, though humans positioned sensors for the data used in this thesis.

The surface of Walker's Mill is almost entirely sandstone with negligible amounts of flora and a lizard or two. The surfaces are not weathered, and the fractured nature of the rock is evident in the lack of uniform floor or ceiling. The cave is damp and most of the walls are covered in a thin layer of water. The surface geometric nature of this cave is not common, particularly of sandstone caves. However, it is a good example of single-material environments and their prevalence.

5.2.1 Variegation

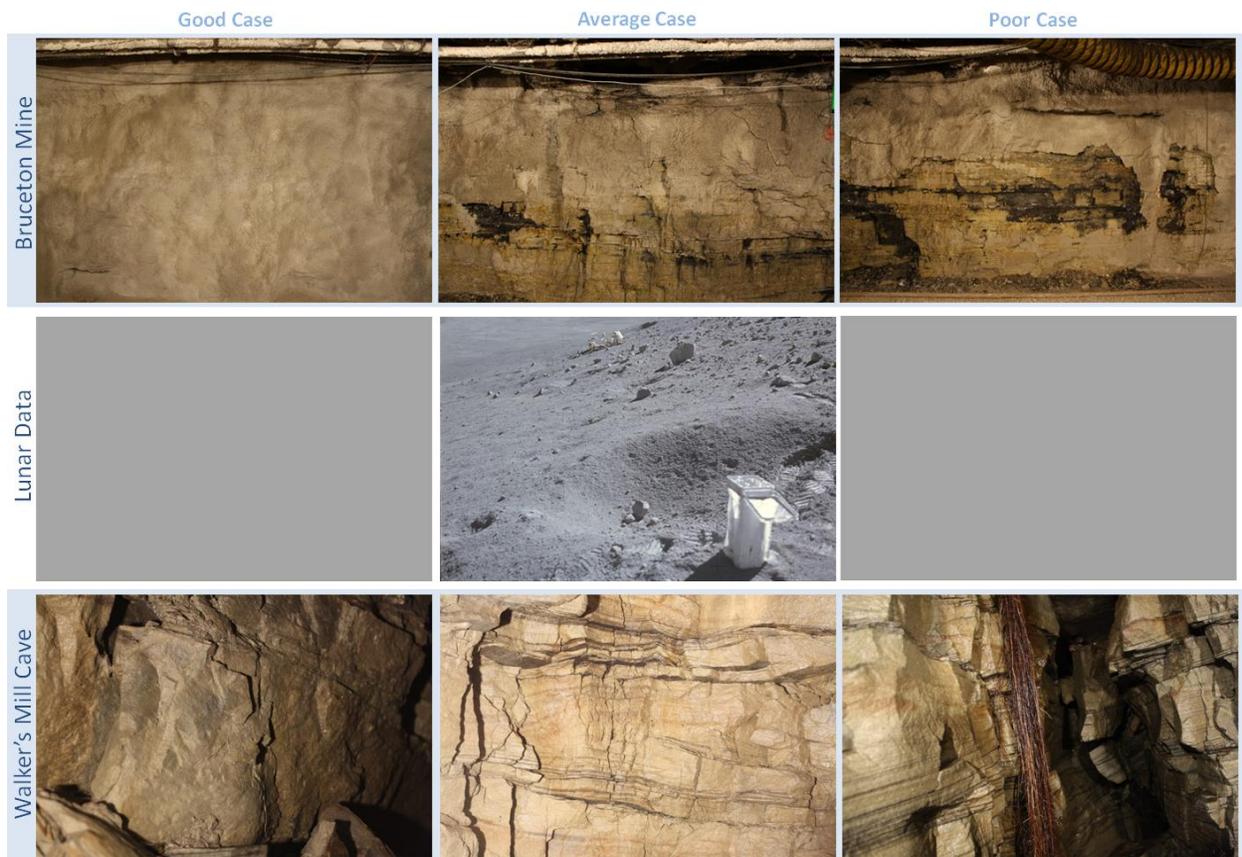


Figure 37. Variegation of Materials in Environments. The walls of Bruceton Mine (top row) show variegation between gunite and coal, and occasional artificial objects. Lunar terrain (center) was only numerically analyzed for the moonyard, which is an artificial average case. Variegation is believed to be very low. Walker's Mill Cave (bottom) shows anisotropic banding of the sandstone material, but variegation of material is very low. The right most image shows a single example of plant roots.

Many surfaces documented in these environments are variegated, meaning that they consist of discrete “plotches” of differing materials. These surfaces are not well represented by a single macroscopic material or albedo, and they are best considered as a combination of surfaces which are each single-material. Vision algorithms utilizing material information should allow for and be capable of distinguishing material changes (see section 2.3). As the planetary *rocky* assumption provides for uniformity of materials (in addition to diffuseness), Lumenenhancement techniques presented here cannot apply to highly variegated surfaces. However, many times variegation is minor and a simple vision solution may mitigate or tolerate some error in these cases.



Figure 38. Variagation Estimation from Images. A color image taken under near-ambient illumination (left) is clustered with the Mean-shift algorithm using high noise tolerance and the L*a*b*. A few large clusters are labeled annotated (right). The number and size of the resulting clusters determines the variegation of the scene.

Variagation in the representative environments was studied by collecting and analyzing imagery. Figure 37 shows examples of these images collected. These images are taken under quasi-ambient illumination to reduce cast shadows and promote uniformity of appearance for the same materials²². These images are then transformed into the L*a*b* colorspace which heuristically separates material from lighting changes, and is particularly effective in diffuse environments. With the luminance removed, mean-shift clustering [with EDISON; Christoudias, et al. 2002] is performed, using an edge tolerance of 0.9 for high noise reduction. Statistics over the number and size of the clusters determine the variegation (Figure 38). A square-weighted *uniformity score* is utilized to give higher weight to a few large clusters as opposed to many small ones:

$$score_{uniformity} = \sqrt{\frac{1}{C} \sum_{i \in C} S_i^2} \quad (5.19)$$

where C is the number of clusters and S_i is the fractional size of the i th cluster (i.e. num_cluster_pixels / total_image_pixels). A higher score represents a more uniform, less variegated scene. It is noted that this technique is neither a classifier - it does not assign material labels - nor is it discriminative – different materials are occasionally grouped together when cluster size is small or edge strength is weak. The uniformity score is only an estimate of the nature and frequency of variegation that has demonstrated utility in this work.

²² Similar to a Lambertian white-out condition

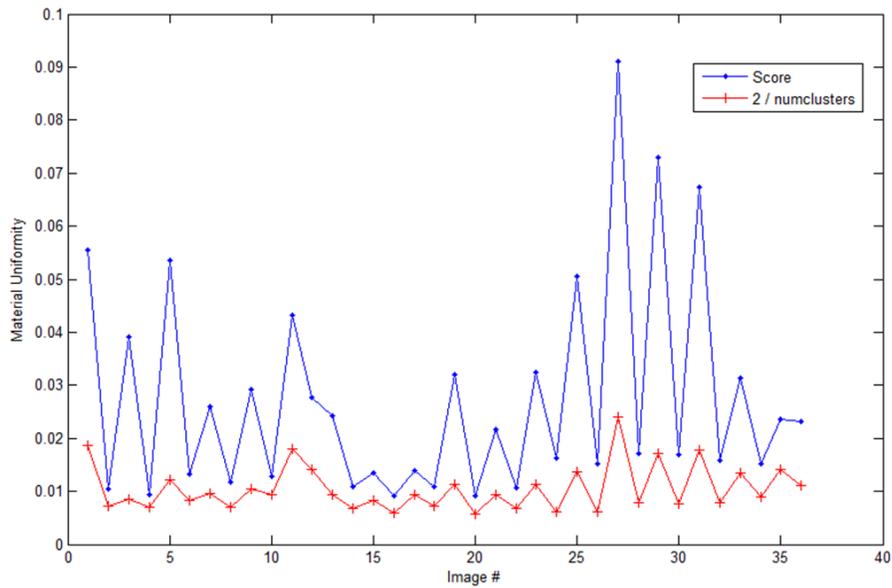


Figure 39. Variiegation in Bruceton Mine Images. The uniformity score (blue) and the scaled, inverse number of clusters (red) are plotted on the same axes for comparison. Higher numbers represent more uniform scenes.

Thirty six images of various walls in Bruceton were collected. The ceilings and ground were not studied as they do not satisfy the domain assumptions and are uninteresting to robot modeling respectively (see above). The statistics of the Bruceton data are summarized in Figure 39. An average of 217 clusters was required per image, with a wide standard deviation of 72 and a mean score of 0.027. Twelve images of Walker’s Mill cave were collected of the walls, ceiling and floor. These were mostly homogenous requiring 160 clusters average with a standard deviation of 30. However, the resulting uniformity score of 0.020 was slightly lower than Bruceton. This is believed to be from the macroscopic color banding of the sandstone, despite known uniformity of material. These bands may share a similar reflectance function, but are of noticeably different albedo. Lastly, only a single overhead image was taken of the moonyard, mostly for comparison with the other environments. The moonyard, a very uniform terrain, and had a uniformity score of 0.18, about 6 times higher than the mine. The terrain was represented with 28 clusters. Figure 40, below illustrates the variegation of an example indoor scene featuring many simple surfaces to be about twice that of Bruceton mine.



Figure 40. Variagation of Indoor Scene. This garage scene required 367 clusters, and produced a uniformity score of 0.0096, less than half that of Bruceton mine.

5.3 Geometric Characterization

Geometry is the final key in the three-part appearance model for understanding environments. This section explores distributions of geometry – both surface and macroscopic – in planetary environments, the suitability of planetary assumptions in explaining these distributions, and the accuracy of surmising these distributions with robotic sensors. Experimentation utilizes data from the exemplary environments discussed in the previous section.

5.3.1 Barrenness

The *barren* property of environments constrains the local smoothness of surfaces. More explicitly, it states that the surface normal, which controls surface irradiance, is related to and can be estimated from discrete surface geometry through differentiation ($\hat{n} = \nabla x$). The amount of discreteness tolerable is of great interest to this work as it defines how densely and how accurately the underlying models must be acquired with a range sensor. Similarly, recall that intensity values in planetary constrained images are directly related to the gradients of surface geometry ($E \propto \nabla x$). It is also pertinent to characterize how well pixel intensity values explain the true surface gradient.

5.3.1.1 Correlation of Image Intensity to Surfaces

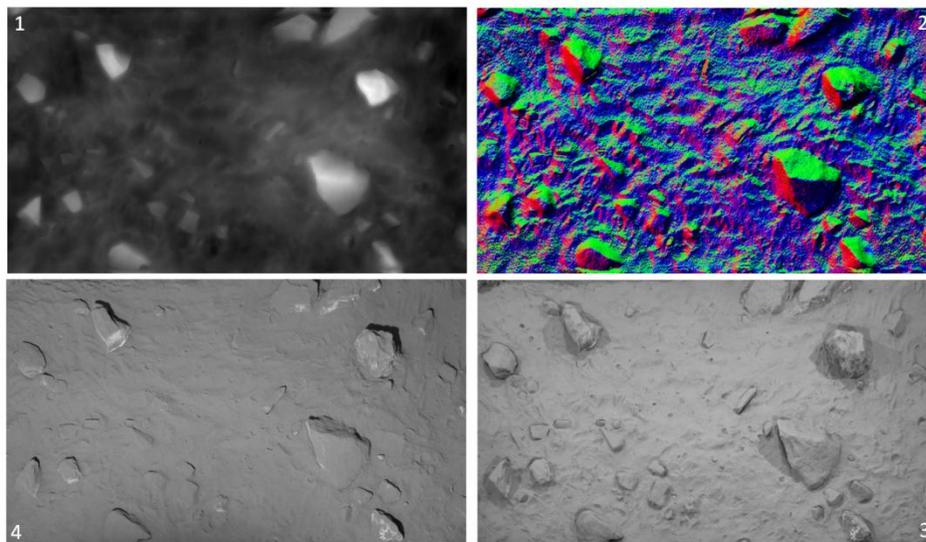


Figure 41. Correlation of Image Intensity to Surface Normals. An ultra-accurate geometric model (1) is used to produce surface normal estimates at each voxel location via differentiation (2). The slant angles calculated from these surface normals are correlated with pixel intensity from an ambient image (3) and a point-illuminated image (4) that satisfies the planetary assumptions.

To determine the correlation of intensity images to surfaces, a high resolution HDR image is taken of the surface under the constrained illumination conditions. This image is registered with the geometry image of a maximum survey-quality scan (1mm range accuracy, 2mm sample spacing) of the scene from the same perspective. This geometry image is a rectangular 2D parameterization of the scene much like a range image. However, the $(N \times M \times 3)$ dimensions explicitly store a cartesian coordinate for each sample, enabling higher accuracy than pure depth parameterizations. After registration, both geometry and intensity images are voxelized to the same resolution (about 2mm spacing). The pixel values are then correlated, using the Pearson product-moment, with the \sin^{23} of surface slant angles which are found using the following equation:

$$\theta = \text{atan} \left(\sqrt{\left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2} \right) \quad (5.20)$$

where $\frac{dz}{dx}$ and $\frac{dz}{dy}$ are the surface gradients. The image is then blurred by half the bandwidth using a Gaussian kernel to average intensity values and the process is repeated. Averaging pixel values determines what amount of image information is noise and how many image samples are needed to optimally measure a surface normal.

²³ The slant angle is complementary to the incident angle.

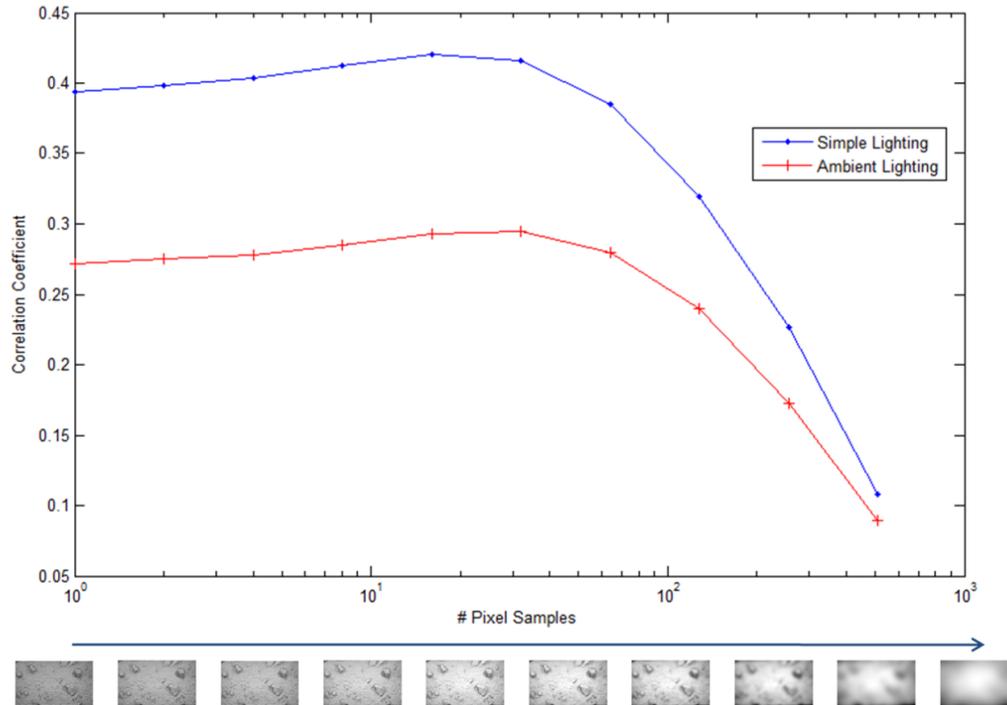


Figure 42. Correlation of intensity values to surface gradients for the moonyard under two different lighting assumptions. This graph shows that simple point illumination promotes superior surface recovery. The optimal value occurs at $n=32$ for ambient and $n=16$ for simple lighting.

This analysis was conducted on the moonyard terrain using both camera-centered point illumination and ambient illumination simulated by many simultaneous sources (see Figure 41). Figure 42 shows the correlation curve for the moonyard over image scales $[1, 2, \dots, 512]^{24}$. The correlation for the point source image goes from $p=0.39$ for raw pixels, to a high of $p=0.43$ at $n=16$, and finally drops to $p=0.12$ at $n=512$. The correlation for images is moderate in absolute terms, but very high for an image taken in environmental conditions. The curve for the ambiently illuminated image, which correlates significantly less, supports the hypothesis that simple, point illumination is required for optimal recovery of geometry from images.

²⁴ A value of 2 denotes a 1/2 scale image or roughly speaking, an average of $(2)^2 = 4$ pixels.

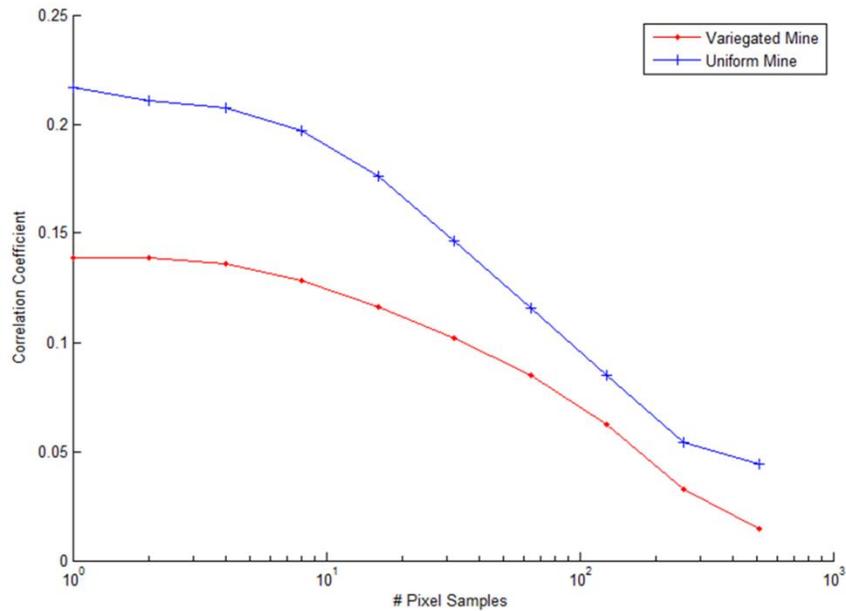


Figure 43. Correlation of intensity values to surface gradients in Bruce-ton Mine Imagery. A variegated scene with patches of gunite and coal is compared against a uniform scene consisting only of gunite. Uniform scenes improve the accuracy of geometric recovery with intensity images.

Analysis was also conducted for Bruce-ton data on representative images (see section 5.2.1 above) to compare variegated (many patches of gunite and coal in the scene) and uniform (gunite only) sections of the mine. Both images were taken under point illumination. Figure 43 shows the correlation curves for the Bruce-ton images. The uniform image continuously drops from a high correlation of $p=0.22$ at $n=1$ to a low of $p=0.07$ at $n=512$. The variegated image follows a similar path, but ranges from $p=0.13$ to $p=0.015$. The uniform Bruce-ton image correlates about half as much as the moonyard data. Curiously, there is no initial improvement in correlation when neighboring pixel values are averaged. This could be the result of the rough gunite surfaces compared to the smooth surface of the dusted regolith.

This analysis could not be conducted for Walker’s Mill cave imagery because the high-resolution survey scanner used to produce clean surface normal estimates could not fit in the entrance (see section 5.3.2).

5.3.1.2 Smoothness with Autocorrelation of Surface Geometry

Local smoothness determines the extent of surface deviation between two sampled points. The local smoothness of planetary surfaces is tested with the same approach as the correlating pixel intensity, except an autocorrelation is preformed with downsampled data instead. Bilinear interpolation in range image space enables comparison of the downsampled and native density models at high resolution. Surface slant estimates are generated from each using numerical differentiation.

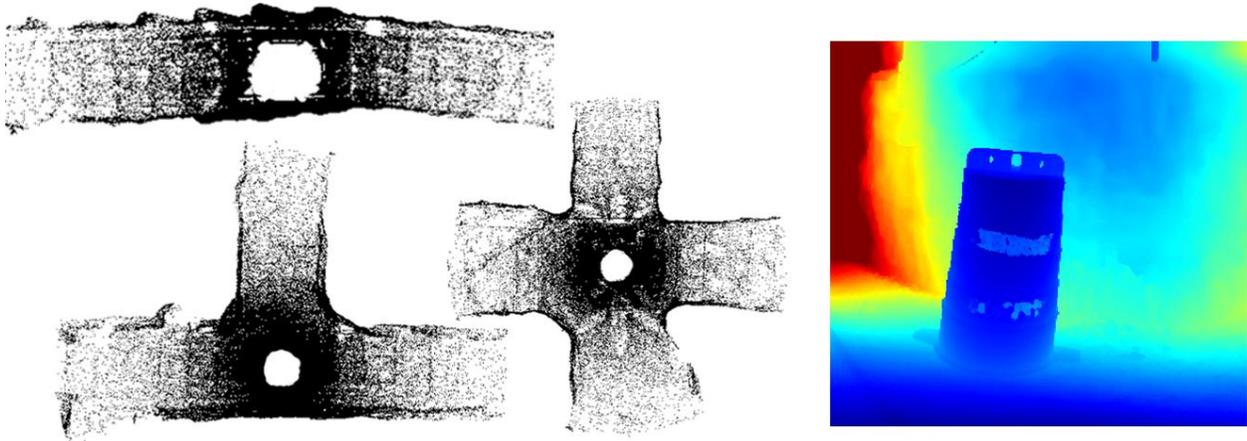


Figure 44. Archetypical Component Geometry in Coal Mines. Point clouds of corridors, 3-ways and 4-way intersections are utilized in testing how macroscopic geometry affects the smoothness constraint (left). The holes in the point cloud are from sensor self-occlusion, and do not effect computation. A barrel against the backdrop of a smooth wall creates an occlusion edge which strongly breaks the smoothness assumption (right).

Three high resolution scans of Bruceeton were collected at archetypical corridor, 3-way intersection and 4-way intersection locations (Figure 44). Each of these scans has a minimum range of 1m and a maximum range of 7m where the areal density of points is 1 per cm^2 .

These component environments are utilized to test *barrenness* as a function of local differentiability as well as macroscopic differences in the number and magnitude of known occlusion edges (i.e. corners) in the scene. Occlusion edges, resulting from range discontinuities create problems for estimation of surfaces (and their normals) from LIDAR and image data. It is not always possible to differentiate between an occlusion and a high gradient surface from a single view, particularly when interpolation or rescaling in image space is performed. Moreover, occluding objects cast shadows, which violate simple source constraints in intensity images.

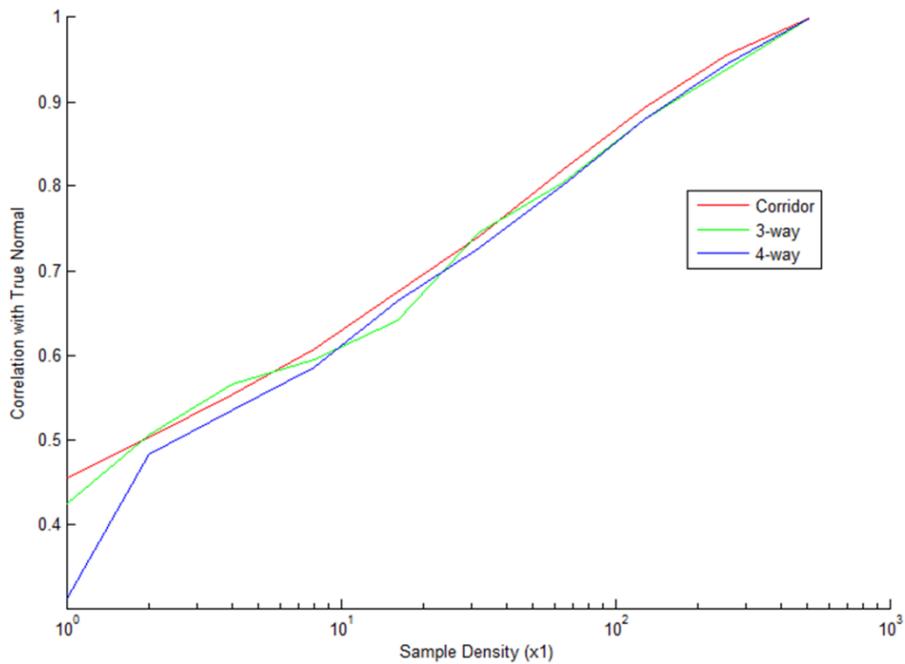


Figure 45. Autocorrelation of Interpolated Normals from Mine geometry. Areal density at the maximum range of 7m is about 4 range samples per cm^2 at native resolution (512x) down to about 0.18 samples per cm^2 for 1x density. Data from three geometrically distinct scenes within the mine are shown: a straight corridor, a 3-way intersection and a 4-way intersection.

Figure 45 illustrates autocorrelation of downsampled range data with its native resolution counterpart. Each sample density uses an average correlation from 10 trials of random downsampling. The low end of the spectrum ranges from 1x sample density to 512x at the high end (which is the native resolution). The graph shows that interpolation of range readings to estimate true surface normals is surprisingly robust. Correlation remains strong even with 8x downsampling. The curves from the different scenes are also very similar indicating that the overall quantity of occlusion edges in each scene is similar and low. Their effect on total correlation is minimal even as their individual values may be way off. Only when sampling density is severely deficient, do the differences in macroscopic geometry play a significant role.

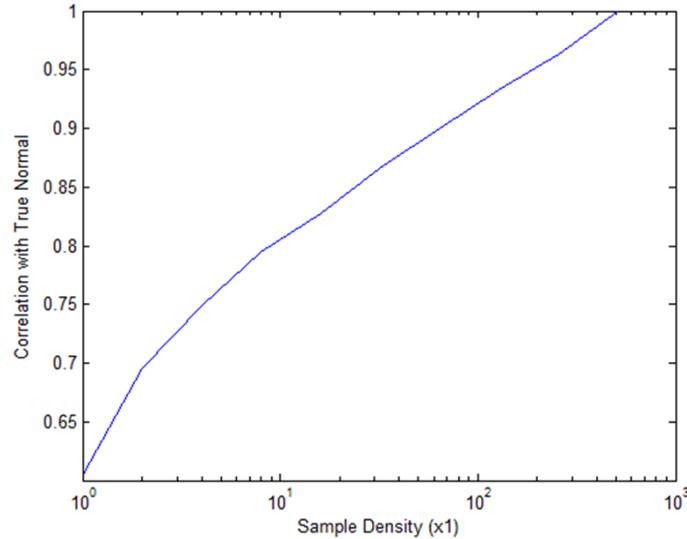


Figure 46. Autocorrelation of Interpolated Normals from Moonyard geometry. Areal density ranges from 1mm grid size at 512x resolution to a 23mm grid size at 1x.

The analysis was also conducted with moonyard data using a surface grid size of 1mm at native resolution. Due to the manageable nature of the terrain, samples are uniform throughout. This presents an opportunity to explore “clean” data featuring less bias from a disparity of ranges and errors like mixed pixels from glancing LIDAR measurement. As such, it represents closer estimation of true surface smoothness, while the Bruceton data represents the aggregate effects of sampling in situ with non-ideal measurement.

Figure 46 shows that the absolute values of moonyard autocorrelation are greater across the entire spectrum of sampling densities. This is within expectation given the features of the lunar terrain. Interestingly, the curve shows very similar smoothness behavior to the mine.

5.3.2 Accuracy of Modeling

Geometry in environments cannot be known to arbitrary precision. Surface geometry is only resolvable to sensing resolutions. LIDAR is perhaps the only way to directly measure geometry; yet it is sparse, subject to noise and exhibits a multitude of nonlinear effects. While sensors are not attributes of the environment, it is important to understand their performance, particularly when the objective is modeling by measuring surface geometry. Sensor characterization informs the intrinsic accuracy of range models gathered with robots as well as contrasts the accuracy and density of different sensing modalities. This information can be utilized in the design and targeting of multi-sensor techniques. A

brief study of sensor performance is included here as part of more expansive environmental modeling work of the author [Wong, et al. 2011].

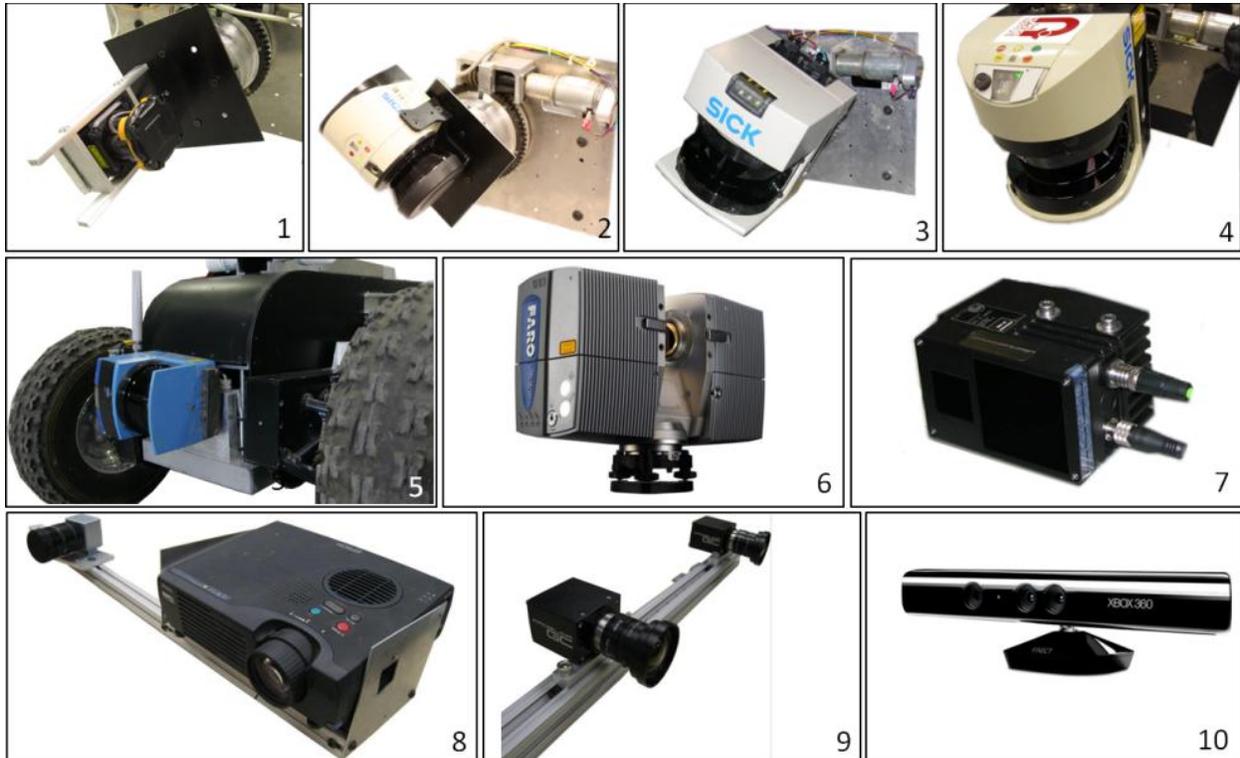


Figure 47. Illustration of Sensors and Configurations Evaluated – (1) rotating Hokuyo UTM-30LX, (2) rotating SICK LMS111-10100, (3) rotating SICK LMS291-S14, (4) rotating SICK LMS511-10100, (5) rotating SICK LMS200-30106 affixed on a mobile robot, (6) Faro Photon80, (7) IFM O3D 201, (8) custom structured light sensor, (9) custom stereo vision sensor and (10) Microsoft Kinect.

The modeling performance of 10 range sensors was evaluated (Figure 47). Sensors were selected based on prevalence in robotics modeling usage and availability. Experimental configurations (i.e. actuation, physical parameters, and external illumination) were chosen to reflect optimality for modeling at a critical sensing distance of 2-8 meters [Omohundro 2007] (Table 6). This study is not intended to be a comprehensive sampling of sensor configuration parameters, but rather a broad sampling of sensor types applicable to planetary spaces²⁵. 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.

²⁵ Applicable in a phenomenological sense: stereo is not applicable in absolute darkness without external illumination and visible structured light is likewise not applicable in daylight. Attributes such as space worthiness are not considered.

Table 6. Evaluated Sensors and Technologies

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, 5mm, 24m clipped

*denotes an in-house implementation

Characterization involved scanning a 1.25m x 1.25m, tiled and colored 3D checkerboard in a controlled laboratory setting. 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 accuracy of most contemporary range sensor technologies.

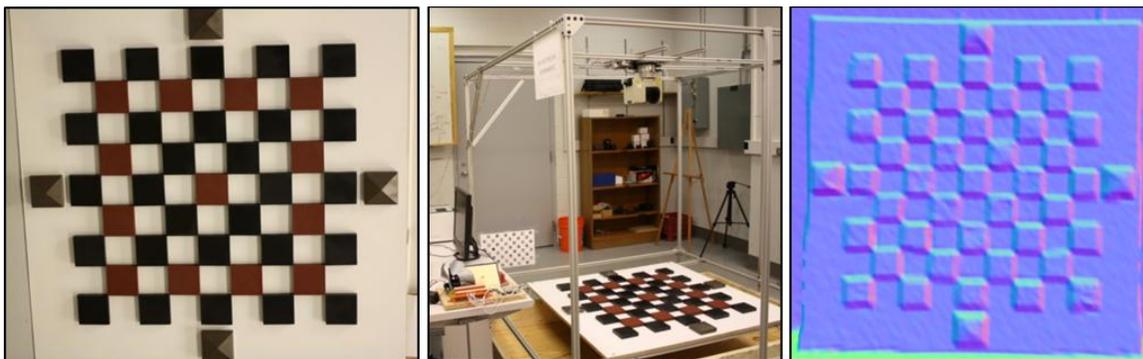


Figure 48. A 3D checkerboard target used for ideal characterization (left), example experimental setup (LMS291 shown) for scanning the checkerboard (middle), and a mesh model of checkerboard generated using range data (right).

Features of the checkerboard are illustrated in Figure 48. 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 used to automate the process of aligning scans as well as testing 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 (Figure 48 - center).

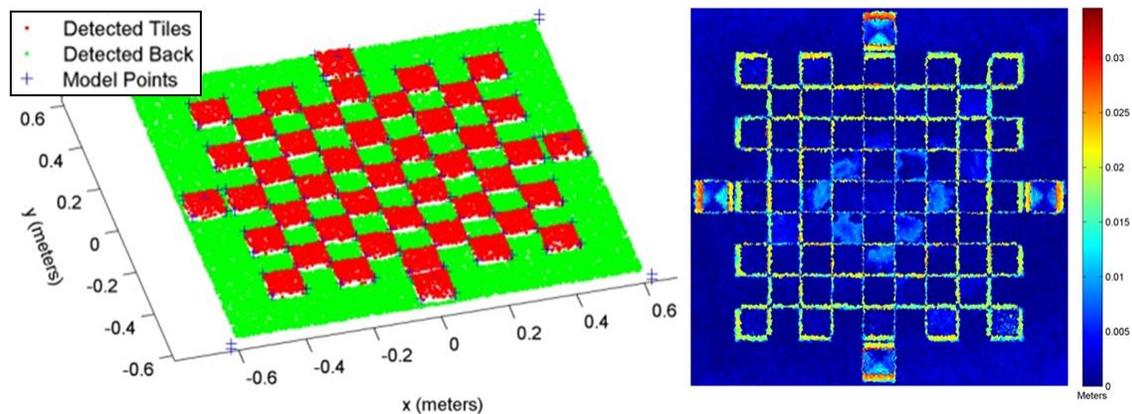


Figure 49. ICP aligned range data from Photon80 with detected tiles and background (left) and range error plot illustrating the “mixed pixel” effect near the edge of the tiles (right).

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 (Figure 49). Points detected as being on 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. Two statistical measures of quality are then computed:

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 here 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 L_2 -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.

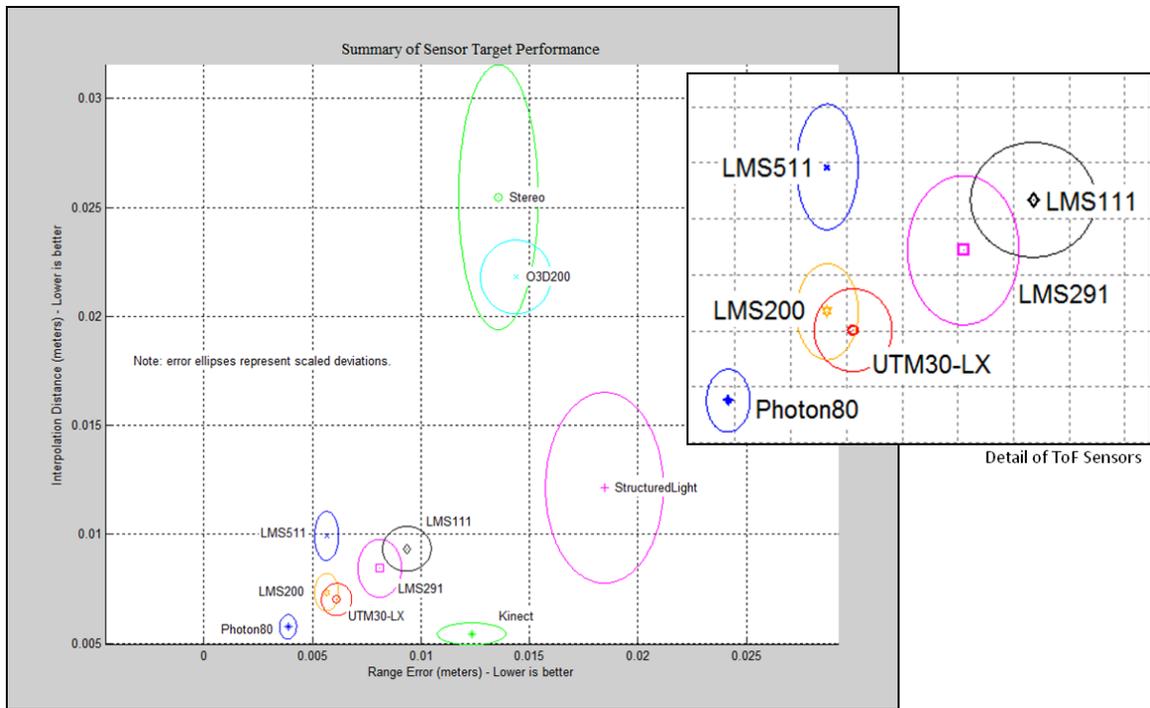


Figure 50. Summary of Ideal Target Characterization with detail of planar time-of-flight LIDAR performance (inset). All statistics are for single-shot measurement where applicable.

Using the metrics described, analysis was performed on sensor scans of the ideal target. The results are shown in Figure 50. The x-axis (range error) is the empirical value of the accuracy and the y-axis (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 as 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. The Photon80 is utilized for collecting all characterization and “ground truth” models in this work. The sensor has a range accuracy of 4mm for a single shot. Multiple returns can be averaged for a more accurate reading, which is the mode of operation utilized. Ground truth models have a three sigma uncertainty of about 2mm. Maximum areal density at the critical sensing ranges of 2-8m are well within 1mm^2 .

All five planar time-of-flight sensors characterized exhibited similar performance in a class below the Photon80, which is consistent with manufacturer specification and intended application. 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. The LMS200 is utilized on the *Cavecrawler* mobile robot featured in this work and has a practical range accuracy of 6mm. 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. The structured light configuration is utilized in the case studies section of this thesis to implement a new hybrid sensor to improve these shortcomings. 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 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 pricepoint²⁶.

²⁶ The Kinect is likely not applicable to field environments due to packaging. However, it is included here for comparison due to popularity in indoor robotics.

5.4 Discussion

Utilizing statistics from material, geometric and environmental characterization, informed comparisons of environments can be made. Table 7 below compares the mine, moonyard and cave using the *barren* and *diffuse* assumptions of planetary appearance and a third attribute that describes uniformity of albedo. Statistics from an indoor garage scene are included as well. These metrics show that the experimental environments are more diffuse, equally barren and highly uniform compared to the indoor scene. Of course, the indoor scene is a single example; however it also represents a highly constrained environment apart where simple vision techniques are applicable and successful²⁷. Thus, these planetary examples compare favorably for the use of imaging.

Table 7. Comparison of Environmental Attributes Satisfying Planetary Appearance Constraints. The indoor garage is a single example intended for comparison. A picture of the scene is shown in Figure 40. Diffuseness value is the representability of dominant materials as Lambertian. Barrenness is the correlation of image intensity with range gradients under point illumination. The uniformity score defined in the prior sections. All metrics are in the range [0,1].

	Diffuseness	Barrenness	Uniformity
Bruceston Mine	0.78	0.22	0.027
CMU Moonyard	0.96	0.39	0.180
Walker's Mill Cave	0.83	n/a	0.020
Indoor Garage	0.50	0.33	0.009

The diffuseness scores in Table 7 are generated by using the occurrence ratios in Figure 31 and multiplying each material by the difference in correlation between the Lambertian and the best-fit BRDF model found with gonireflectometry. The scores are given by the following calculations:

Bruceston

$$(45\% \text{ gunite} \times 0.82) + (15\% \text{ coal} \times 0.87) + (25\% \text{ gravel/dirt} \times 0.82) + (15\% \text{ other} \times 0.5)$$

Moonyard

$$(90\% \text{ regolith} \times 0.96) + (10\% \text{ limestone} \times 0.98)$$

Walker's Mill

$$(99\% \text{ sandstone} \times 0.84) + (1\% \text{ other} \times 0)$$

The diffuseness of gravel and dirt are approximated from limestone and oren nayer soil material. Diffuseness of the "other" materials are from conservative estimates. The indoor scene is approximated

²⁷ The scene is a part of the "highbay" at CMU, academic laboratory environment where robotic vision systems are frequently (and successfully) tested.

from the qualitative prevalence of semi-gloss flooring, windows, and other metallics in comparison to Lambertian walls and cardboard boxes.

The barrenness score is estimated from the correlation of image intensity values to surface slant at highest resolution and under point illumination. The value for the indoor scene is for ambient illumination as it was impossible to light the scene with a point source. Thus, it is not an entirely accurate comparison. Barrenness of the cave was not estimated because of lack of high resolution geometric data. However, it is likely that this environment is less barren than the mine due to the jagged interior surfaces. Lastly, the uniformity scores are directly is from the variegation clustering experiment.

Limitations

The analysis here is intended to only guide and inform about the validity of the planetary constraints in example environments. The findings are only directly applicable to these specific datasets in the experimental environments. This data is subject to noise, sample bias, and gross variability which cannot be accurately estimated in such continuous spaces. Blowing dust, dripping water, and other ephemeral effects are all artifacts of sensing in the field which ultimately cannot be ignored. These environments are moreover not wholly representative of what may be considered “planetary” in common use. The aforementioned Martian daylight surface diverges significantly from these examples; however, the nighttime may be sufficiently similar. Volcano fields may appear more rocky and barren than environments analyzed here, but scintillate at specific illumination and view angles. Discretion and probabilistic appreciation must still be used in applying inferred properties of the domain to specific applications and environments.

Chapter 6:

Planetary Case Studies for Lumenhancement

This section presents application-oriented case studies of Lumenhancement. These are approaches designed to solve specific problems using planetary domain knowledge, but are applicable to many other environments through generalization of appearance. These studies also document the extensive field validation of the thesis and implementation on working robot systems.

6.1 Camera and LIDAR Fusion for Super-Resolution Modeling

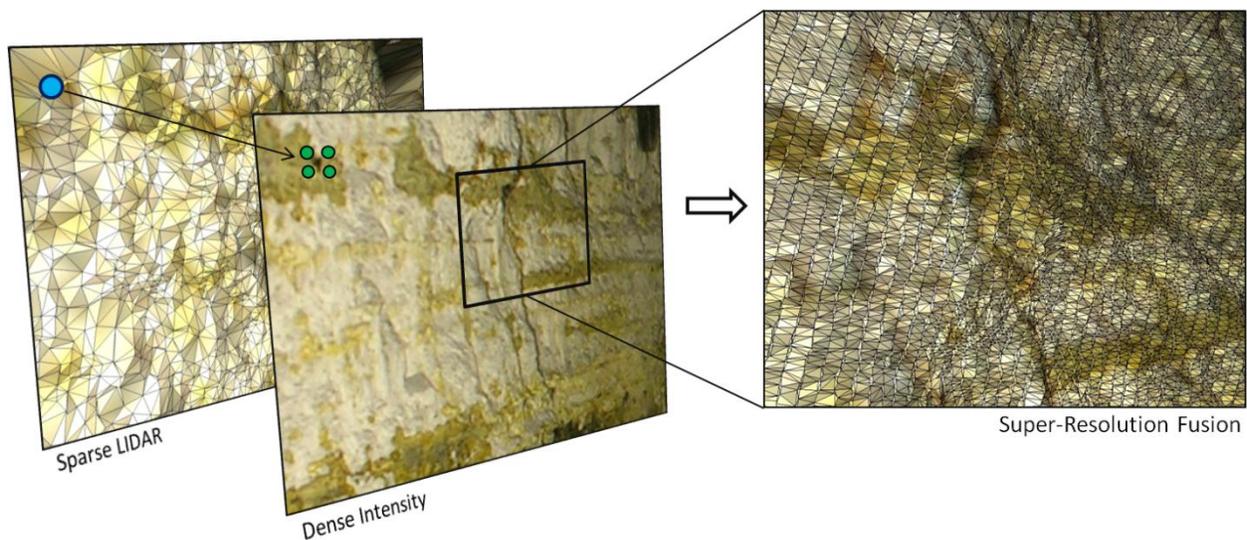


Figure 51. Sparse LIDAR range readings(left) and high resolution intensity images (center) can be fused to create super-resolution models (right) with a Markov Random Field.

Mine accidents including those at Quecreek, Sago and Crandall Canyon highlight the urgency of estimating accurate 3D geometry in mines. Systems have been employed to map mines, from virtual reality systems for training rescue personnel [Boulanger, et al. 2001] to automated survey robots and post accident investigation [Morris, et al. 2006]. While many of these systems use state-of-the-art direct range measurement sensors, LIDAR sensors alone cannot meet the resolution, size, power or speed requirements to produce quality mine maps in a practical amount of time.

Absolute range sensor data can be fused with high-resolution CCD imagery to achieve a quantitative increase in range data accuracy and density in a process called *super-resolution*. While this application of Lumenhancement targets artificial subterranean voids, the technique is applicable to any similar planetary spaces where assumptions can constrain the image formation problem. As both color and geometric information are of interest, cameras and range sensors commonly exist on modeling platforms [Morris, et al. 2006]. Thus, integration of the method presented here requires only calibration and low processing overhead.



Figure 52. A mine inspection robot with active illumination, *CaveCrawler*, in process of corridor modeling (left) and raw fisheye photographic data (right). Note that the left image utilizes a different lighting configuration than presented here.

The results from field experimentation in a working mine are discussed in detail. Dense visualization techniques enabling mesh quality models to be displayed and updated in real-time on GPU hardware are also explored.

6.1.1 Foundational Work

The fusion of range and imaging sensors to improve 3D model quality has been studied in depth [Li 2001; Diebel, et al. 2005; Torres-Mendez 2005; Gould, et al. 2008]. A general model for fusing raw LIDAR and image data into super-resolution range images using a Markov Random Field (MRF) was explored in Diebel and Thrun's seminal paper [Diebel, et al. 2005]. MRFs are undirected graphs that represent dependencies between random variables and have been used extensively in computer vision for noise removal, feature matching, segmentation and inpainting (see [Li 2001]). The popularity of the MRF stems from the ability to model complex processes using only a specification of local interactions, relevance to the regular grid nature of CCD images, and the maximum *a posteriori* (MAP) solution requiring only direct convex optimization in many cases.

Diebel and Thrun surmised that higher resolution intensity (color) data could be used to texture range images and increase the range accuracy of interpolated points. The results in a uniformly and sufficiently illuminated regular office environment are quite compelling. Cameras are able to turn LIDAR scans into dense range images with very low computational overhead. However, the assumption that an image provides relative range information, even locally, is tenuous in unstructured environments. Generating 3D geometry from a general 2D projection is an ill-posed problem. The ability of Diebel's method to smooth point clouds using areas of flat image information was convincingly shown, but the converse of enhancing a point cloud using image texture was not. Recent research in range/camera fusion using MRFs include [Torres-Mendez, et al. 2008; Gould, et al. 2008]; all of which also target indoor application.

This research extends MRF-based super-resolution to subterranean environments such as mines, caves, lava tubes and sanitary pipes. These environments have unknown but slowly varying albedos with a dominant diffuse reflectance term. These naturally-dark, enclosed spaces also require active illumination to image, enabling the use of calibrated lighting. With these assumptions we are able to provide a stronger depth estimate for texturing the interpolated LIDAR data.

6.1.2 Fusion in the Markov Random Field Framework

A range image is used as the common representation for fusion. The 3D range cloud data is registered to the pinhole of the camera, forming a range map (R) via projection of distances onto the $n \times m$ image plane at equivalent resolution. Many pixels in the range map will not contain range measurements; these holes are filled from nearby data through bilinear or nearest neighbor interpolation. The color image data can be then converted to intensity values or used as a raw RGB vector (I). A lattice MRF is

formed where there is a single range and intensity measurement associated with each node (Figure 53). This is similar to the MRF fusion method documented in [Diebel, et al. 2005]; however, the image gradients are instead numerically integrated in this framework.

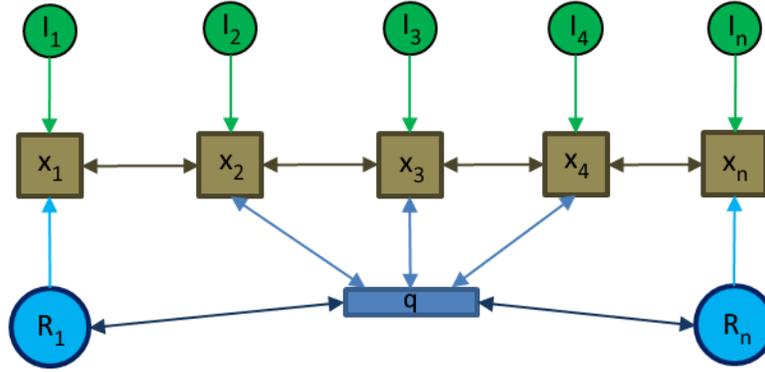


Figure 53. Markov Random Field Graphical Model. Green nodes (I) represent the image pixel data, brown nodes (x) represent the hidden true range value to be estimated, aqua nodes (R) represent the sparse range data and the blue node represents the interpolation uncertainty estimate. There is 1 pixel value for every hidden node (x), but there may be many nodes without a corresponding range value (R).

The range map potential (6.1) promotes agreement between the estimated variables and the interpolated range data.

$$\Psi = w_1 \sum_{i \in L} (R_i - x_i) \quad (6.1)$$

The smoothness prior (6.2) regularizes large changes in the range estimate and like the image potential (6.3) connects potential transfer from a node to its neighbors.

$$\Omega = \beta \sum_{i \in L} \sum_{j \in N(i)} (x_j - x_i)^2 \quad (6.2)$$

$$\Phi = \alpha \sum_{i \in L} \sum_{j \in N(i)} (x_j + \nabla I_{ij} - x_i)^2 \quad (6.3)$$

where relative weights are given by:

$$\begin{aligned} \alpha &= w_2 \exp(-c \cdot \sigma) \\ \beta &= w_3 (1 - \exp(-c \cdot \sigma)) \end{aligned} \quad (6.4)$$

The image gradient is a reasonable predictor of depth change across neighboring pixels. However, integrating the gradient to produce depths over a large locality is prone to drastic shape distortions. The range estimate can be used to regularize numerical integration of the intensity gradient. Moreover, it can ensure that nodes with true range readings are never changed. The weights α and β are relatively scaled by an interpolation distance uncertainty (σ) for some weights w_1 and w_2 (6.4). σ can be generated from the range image during inpainting by using the Matlab command BWDIST, for example. The potential function corresponds to a Gibbs distribution of the form:

$$p(x | R, I, \sigma) = \frac{1}{Z} \exp\left(-\frac{1}{2}(\Psi + \Omega + \Phi)\right) \quad (6.5)$$

$$x_{mle} = \arg \min_x f(\Psi + \Phi + \Omega) \quad (6.6)$$

Solving for the MAP of the distribution requires running a gradient descent algorithm on the target variables x in (6.6), where Z is the partition function [Diebel, et al. 2005].

6.1.2.1 Structure from Shading

The image gradient ∇I_{ij} in (6.3) can apply to either raw pixel data or better estimates of depth from the camera. As scene geometry cannot be ascertained from a single image without assumptions, often no better estimate exists. Definite reconstruction requires knowledge of image formation parameters like light field, surface reflectance (BRDF) and albedos. However, if assumptions like those commonly made in Shape-from-Shading are valid, as in the planetary domain, the amount of certainty is greatly increased.

The illumination and reflectance assumptions are appropriate for subterranean environments. Most dry underground mines and caves are located in Lambertian rock and many coal mine interiors are additionally covered with diffuse material like Shotcrete [Clements, 2003]. Low amounts of metallic meshing, industrial equipment, water and retro-reflectors are present, but the contribution of these specular surfaces can be reduced using the method documented below and in [Mallick, et al. 2005]. Robots in these naturally dark environments can be fitted to carry small area light sources for photography which produce simple light fields.

The MRF image observation (I) is estimated using Shape-from-Shading given the above assumptions. A lightness-based direct normal estimation method which uses range information is given below, but other techniques exist (see background section 4.3). This method factors range information to allow

varying albedos and trades accuracy for feature preservation. The effect of the light source's irradiance fall-off is first removed from the raw image data (E_0). The following irradiance correction model for small area sources is assumed (6.7):

$$E_{unbiased} = \gamma(E_0) \cdot R^n \quad (6.7)$$

The radiometric function (γ) maps pixel values to irradiance, (R) is the interpolated depth estimate and (n) is the irradiance fall-off factor. For ideal point sources $n = 2.0$, while $n < 2.0$ for near-field area sources. The experimental setup described below exhibits an empirical decay of $n = 1.27$. The corrected image ($E_{unbiased}$) is devoid of a near-field illumination intensity bias from the use of an area source. Moreover, the compensation of intensity enables smooth color alignment when stitching several scans together.

Converting RGB color into a single intensity value provides compactness and symmetry, and also minimizes chromaticity effects. Color space transformations such as CieLAB or YCbCr are often used to heuristically isolate the lightness component of an image, discarding chromaticity and albedo. The SUV transformation [Mallick, et al. 2005] describes a class of physics-based specular-invariant color spaces produced by rotating the RGB space such that a single channel is aligned with the illuminant color vectors. This method has produced excellent results with single-source images and enables many Lambertian algorithms to handle a large set of environments with specularities. The specular invariant image, as defined in equation (6.8)-(6.9), is used in experimentation:

$$[s, u, v]^T = R_r(\theta) \cdot [E_{unbiased}^{(r)}, E_{unbiased}^{(g)}, E_{unbiased}^{(b)}]^T \quad (6.8)$$

$$E_{inv} = \sqrt{u^2 + v^2} \quad (6.9)$$

$R_r(\theta)$ is defined as a (3×3) rotation matrix that aligns the red channel of an $\{r, g, b\}$ triple with the source color. The magnitude of the $\{u, v\}$ components is taken to be the diffuse image.

An albedo map is subsequently generated from the diffuse image using Blake's method for lightness computation [Worthington 2005]. Perceived intensity is a multiplicative relationship between surface slant angle and reflectance²⁸. The log image separates these components into additive terms. Scene

²⁸ Recall the Lambertian BRDF: $E = \frac{\rho}{\pi} (\hat{n} \cdot \vec{l}) = \frac{\rho}{\pi} |n| |l| \cos \theta$

albedos can be recovered from the gradient of the log diffuse image by thresholding to remove small changes and integrating. It is noted that the problem can be recast as finding the log albedo map (δ) that minimizes the following equation:

$$\arg \min_{\delta} \left| \frac{\partial}{\partial x} \delta - T_{\sigma} \left(\frac{\partial}{\partial x} \log E_{inv} \right) \right|^2 + \left| \frac{\partial}{\partial y} \delta - T_{\sigma} \left(\frac{\partial}{\partial y} \log E_{inv} \right) \right|^2 \quad (6.10)$$

where (T_{σ}) is the threshold function. Exponentiating (δ) with the proper constant of integration produces the albedo values (6.11). The constant can be estimated from the range data to minimize depth discrepancy in the reconstruction.

$$\rho_{est} = \exp(\delta + c) \quad (6.11)$$

Lastly, surface normal approximations for every pixel are obtained by solving the Lambertian reflectance model:

$$E_{inv} = \rho |n| |l| \cos(\theta_{nl}) \quad (6.12)$$

$$\theta_{nl} = \arccos \left(\frac{E_{inv}}{\rho_{est}} \right) \quad (6.13)$$

The polar estimates (θ_{nl}) are combined with azimuth estimates (ϕ) from the range image. Range data is taken to be a reasonable indicator of the gradient direction while image intensities modulate the gradient magnitude. An integrable surface is constructed from these normals using the method of [Frankot, et al. 1988]. The surface reconstruction is passed into the MRF as a second range image. An accurate surface is neither required nor preferred from this method. Instead, preservation of high frequency detail is preferred, while global consistency is enforced by the decoupled MRF pass.

6.1.2.2 Algorithm Summary

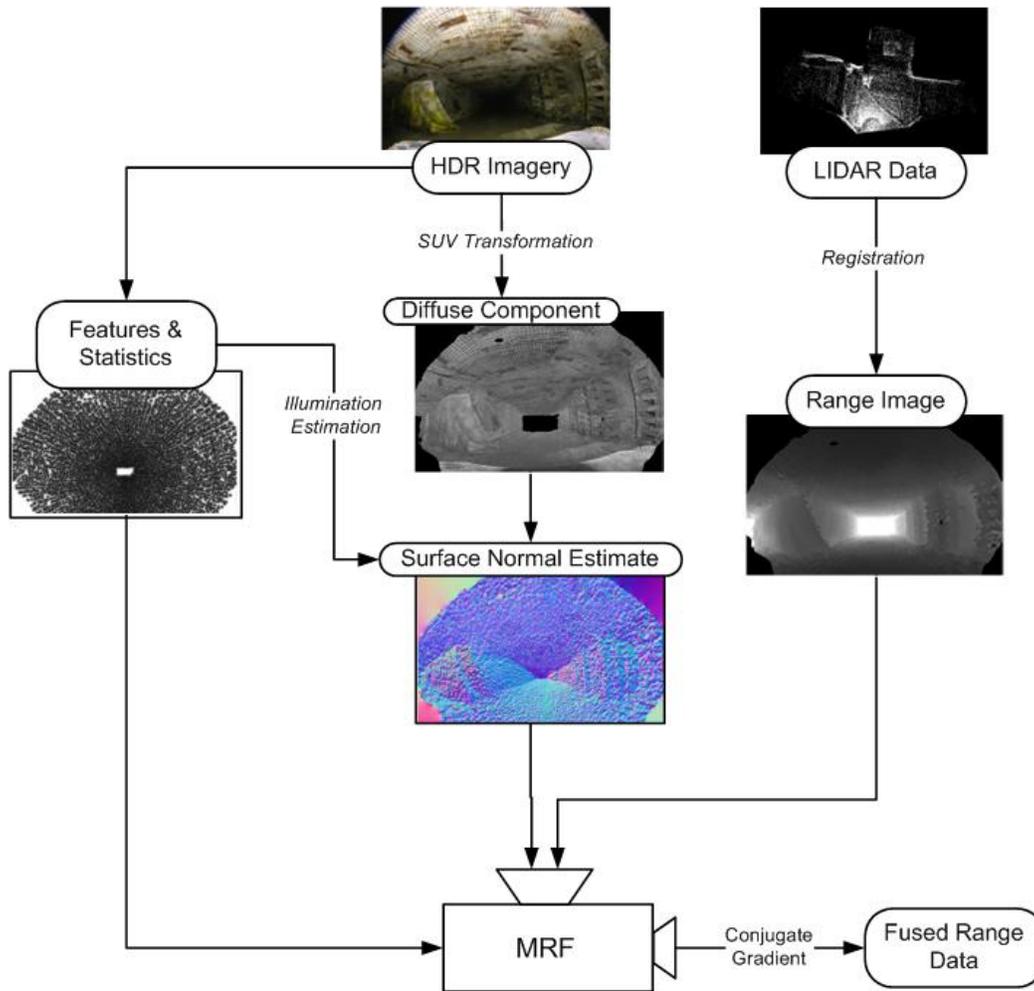


Figure 54. MRF Super Resolution Process. (1) Raw LIDAR point data is converted to a range image from the camera perspective. (2) Specularities are removed from the color HDR imagery to produce a diffuse image. (3) Surface normals are estimated from the diffuse image using shape from shading. (4) The surface normals and the range image are fused in the MRF framework.

A flow chart overview of the technique presented is shown in Figure 54. Raw LIDAR data is first projected into the space of the image, and resampled (interpolated) to form a co-registered range image at the resolution of the color image. Then, HDR color images taken under controlled illumination are transformed into a purely diffuse intensity image using the SUV transformation and knowledge of the spectrum of the light source. The diffuse image along with image features like saturation, illumination and albedo estimates are utilized in a shape-from-shading approach to generate surface normal estimates at every pixel. An MRF fuses the range image, surface normals and uncertainty map into a single high-resolution depth map.

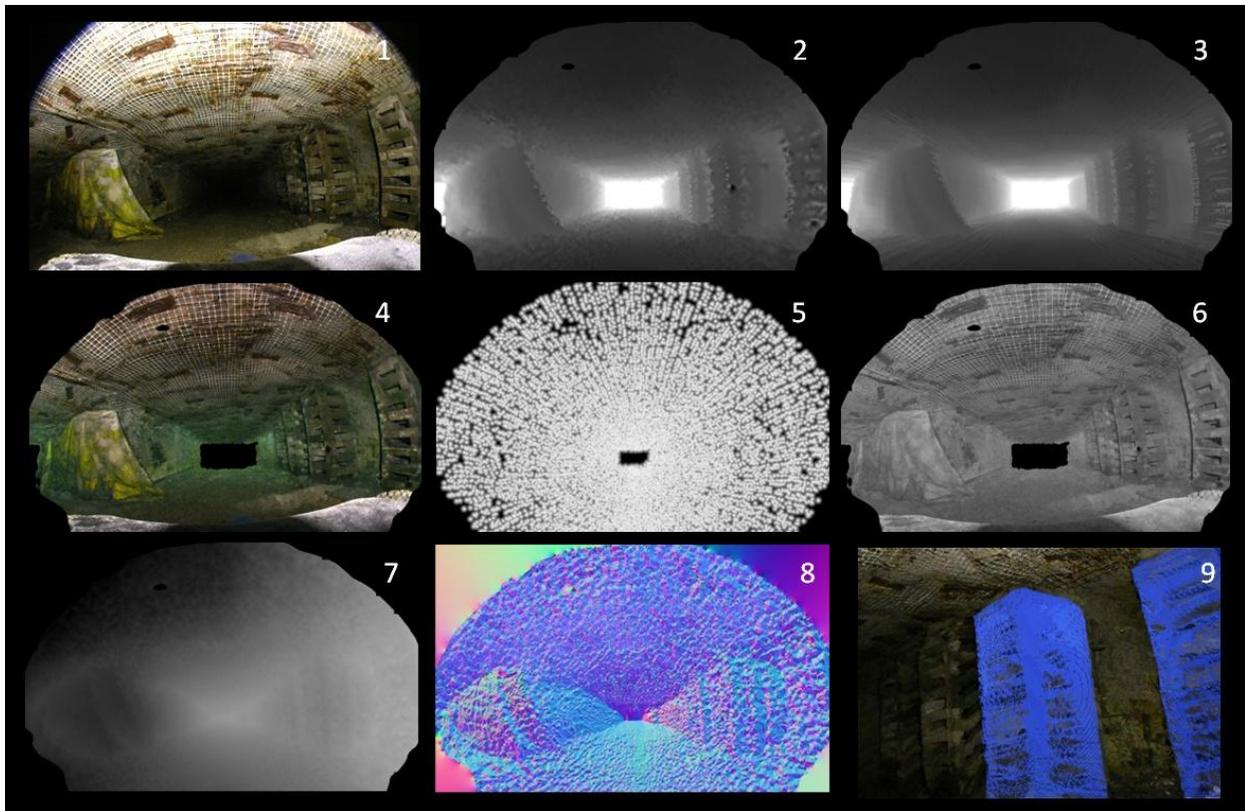


Figure 55. Intermediate Representations of Data for Fusion. (1) Raw fish-eye Image of a mine scene. (2) Range Image (depthmap) from raw LIDAR readings. Depthmap shown is warped to the space of the fish-eye image. (3) Ground truth depthmap. (4) Irradiance compensated color image, clipped to the boundaries of the LIDAR data. (5) Interpolation Uncertainty map. White values indicate scan points, while varying degrees of gray indicate increasing interpolation distance between scan points. (6) Specular-Invariant image after SUV transform. (7) Shading estimate from intensity image. (8) Surface Normal map from shading estimate utilized in MRF. RGB channels correspond to magnitude in XYZ Cartesian coordinates of unit normal vector. (9) Super-Resolution point cloud generated using MRF technique, showing detail of roof supports (right side of image).

Some examples of intermediate data representations in the fusion process are shown in Figure 55. The data is from an underground mine scene with a mine curtain on the left, roof supports on the right and mesh, which stabilizes the exposed rock of the ceiling.

6.1.3 Experimental Results

The experimental setup uses both a continuously rotating planar LIDAR scanner and an 8 megapixel DSLR camera mounted to a mine robot, *CaveCrawler*. A small area light source is also mounted along the same axis to minimize cast shadows in the image. This replaces the normal flood lighting for the imager. The scanner has a practical throughput of $\sim 40,000$ points per second. The points are aligned along concentric rings with 0.5° angular separation in a 180° hemisphere in front of the unit. The camera takes

hemispherical images using a constant angular resolution fisheye lens with a 182° field of view. The sensor mounting configuration and example data are shown in Fig. 1 below.

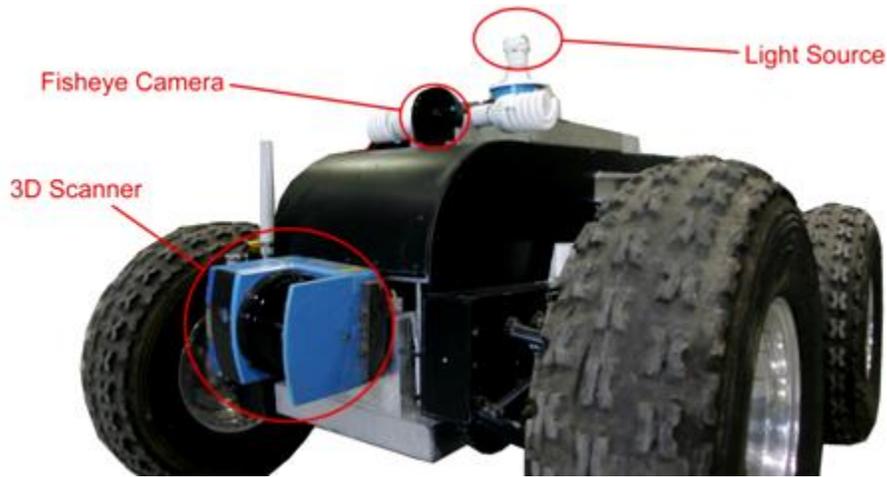


Figure 56. (Left) Experimental setup with 1. LIDAR scanner, 2. Fisheye Camera, 3. Light Source. (Center) Raw fisheye imagery. (Right) Ground truth range image.

Thirty complete datasets consisting of LIDAR scans, High Dynamic Range (HDR) imagery and robot odometry were collected from the Bruce Research Coal Mine in Pittsburgh, PA. LIDAR scans averaged 600,000 points. HDR images were each generated from a series of 5 images corresponding to exposures times of $\{1/4, 1/2, 1, 2, 4\}$ seconds using the method described in [Debevec, et al. 1997]. The 1.0 second exposure image was used as the Low Dynamic Range (LDR) reference image for analysis. An additional 16 datasets of LDR-only imagery were also collected.

A ground truth range map was generated for each LIDAR scan using the full point cloud. Multiple measurements mapping to the same pixel were averaged. The scans were subsequently down-sampled to 25,000 points and interpolated into a range image for testing the method. The datasets were further partitioned into test sets (25 HDR + 16 LDR) and training sets (5 HDR). Optimal weighting factors were learned using a simplex search on the training set, while validation occurred in the test set.

Table 8. Summary of Interpolation Accuracy

Set #	Bilinear	Reference MRF	Proposed	Improvement
1	5.3	5.3	4.8	9.9%
5	3.0	3.0	2.7	12.8%
9	2.9	2.9	2.7	7.6%
13	3.5	3.5	3.0	14.8%
21	7.4	7.4	6.3	17.9%
25	7.6	7.6	6.9	9.3%
32	5.9	5.9	5.2	12.2%
38	10.5	10.5	9.1	15.9%
Total (41)				12.2

*10 selected data sets reproduced here. Rest are omitted for clarity.

**mean per-pixel error in units of centimeters

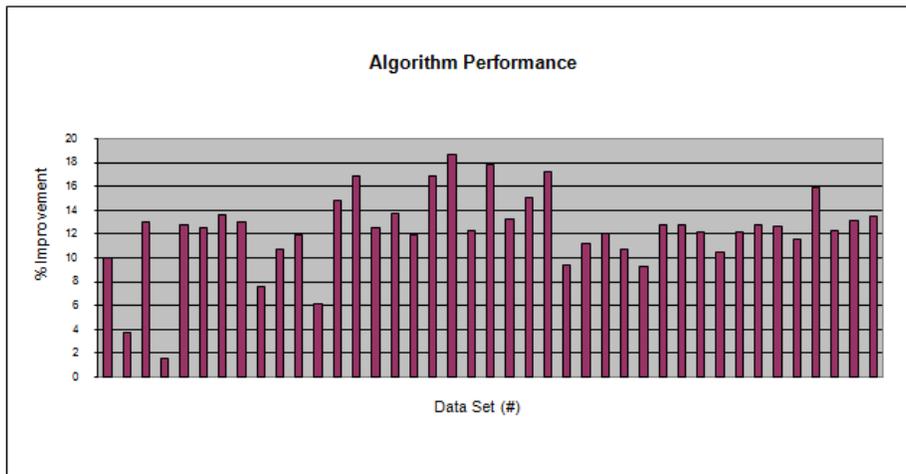


Figure 57. Reconstruction Improvement vs. Raw Interpolation.

The results of the experiment are summarized in Table 8 and Figure 57. The Lumenhancement-inspired method is compared against Diebel’s method and raw interpolation. In this comparison, the mean per-pixel error between the reconstructed range map and the ground truth map is used as the benchmark. Ground truth data points outside the convex hull of LIDAR values in the interpolated map are discarded due to skew in scoring extrapolated points. The usable pixel area is determined for each scan by the number of saturated pixels, the range image convex hull and removal of high-gradient probable error values.

An example reconstruction from a single view point scan utilizing the mine scene described in Figure 55 above is shown in Figure 58 and Figure 59. Detail of features of interest to inspection – the roof supports

and mine curtain – illustrate both a qualitative increase in measurement density and a quantitative increase in range accuracy.

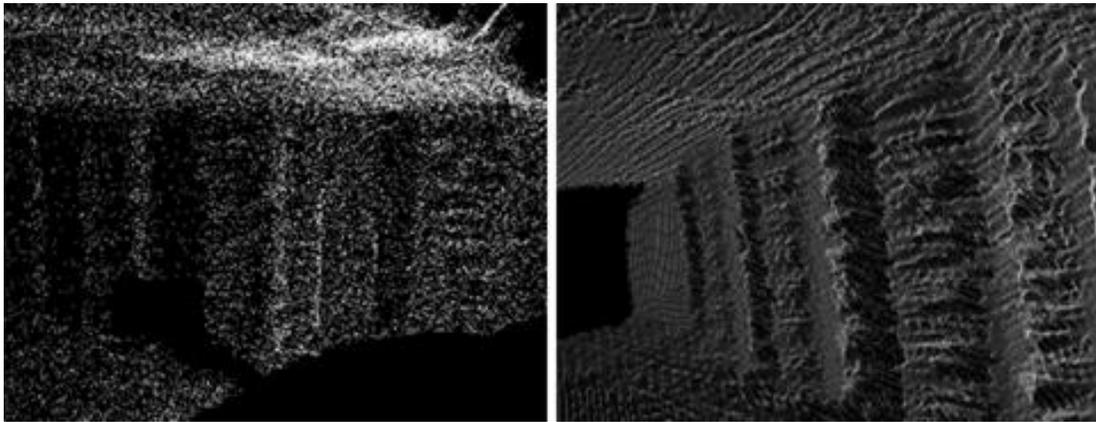


Figure 58. Point Cloud of Cribbing. Low resolution cloud (left) and high resolution reconstruction from algorithm (right) showing stacked timbers supporting the roof.

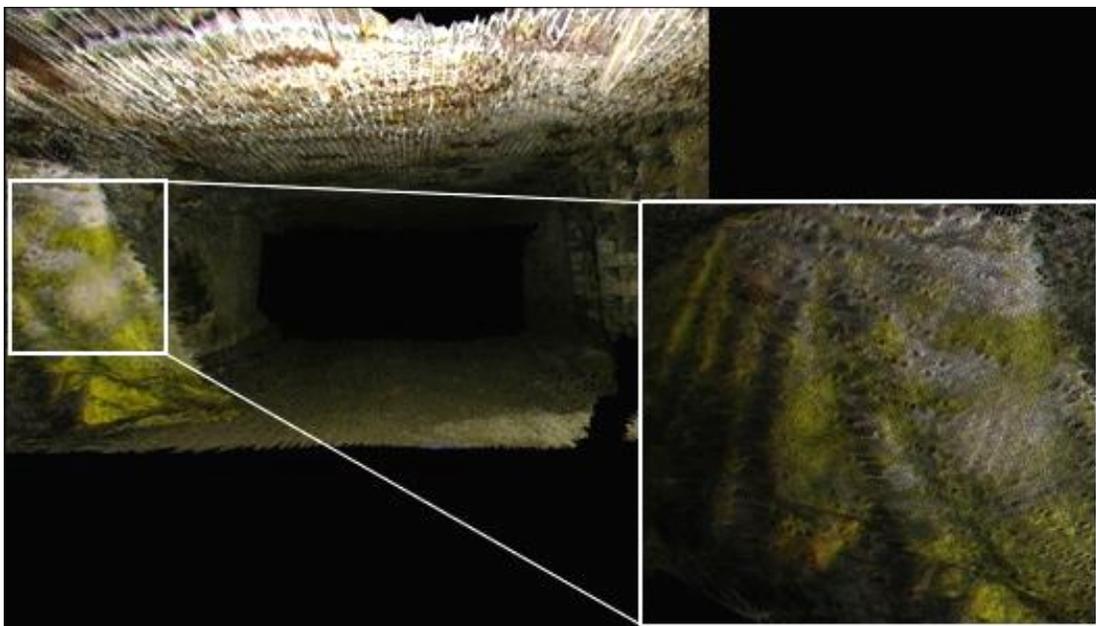


Figure 59. Colorized 3D Reconstruction. Full scene (left) and mine curtain detail (inset and right).

Table 9 summarizes important statistics of the field experimentation. In particular, individual results for the LDR and HDR imagery are given for comparison.

Table 9. Summary of Super-Resolution Experimentation

Quantity	Details
Total Test Datasets	41
{HDR, LDR-only} Datasets	{25, 16}
Interpolation Improvement	
Mean	12.2%
Max, Min	19.2%, 3%
Density Statistics	
LIDAR downsample	25,000 points
Ground Truth LIDAR	669,834 points
Mean Resultant	1,045,358 points
Mean Increase	41.8 x
Image Usability Information	
LDR Saturated	3.17% of total pixels
HDR Saturated	4.20×10^{-2} % of pixels
HDR Accuracy Increase	20.5% over LDR-only
HDR Density Increase	51.5% over LDR-only

Additional data of two corridors were also collected at the Bruceton Mine along evenly spaced intervals roughly 3 meters apart. Using robot odometry and Iterative Closest Point (ICP) alignment, multiple scans were up-sampled using super-resolution Lumenhancement, fused together and color/illumination compensated. These models represent some of densest, most comprehensive mine reconstructions to date using a mobile robot. The results appear below:

Table 10. Corridor Modeling Statistics

Model #	# of Scans	# of Images	# Points
1	4	16	5,543,451
2	8	32	9,680,105



Figure 60. Mine Corridor 3D model. (1) External view. (2) Internal view with rail tracks.

The density of resulting points opens many possibilities for visualizing data. While decimated meshes provide continuous surfaces and fast hardware rendering, large meshes have high pre-computational overhead, correct surface polygonizations are non-trivial and display devices are not optimized as polygons approach pixel size. Points with color and normal information (surfels) are an alternate way of visualizing this information [Pfister, et al. 2000]. With the advent of general pixel shader hardware and high throughput measurement techniques, *point rendering* has become an alternative to meshing, especially when real-time structural updates are necessary.

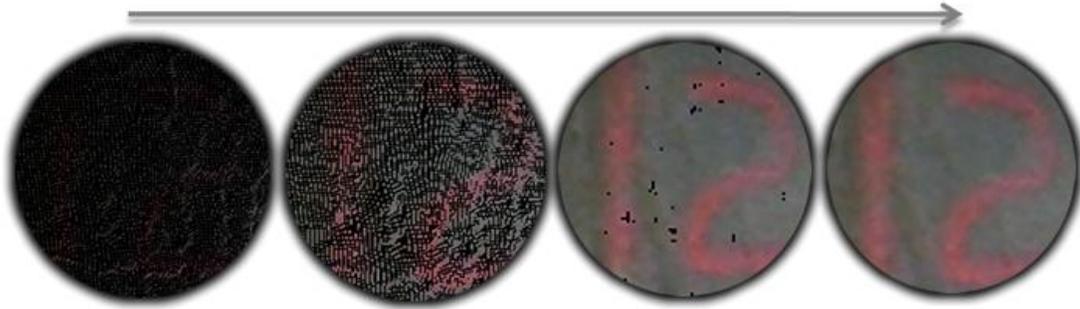


Figure 61. Detail of Hole Filling Process. Dense underlying geometric points are connected with their neighbors using multi-scale texture. Surface normal estimates from image data provide discontinuity checks and perception of shape.

The results are displayed using a hole-filling method similar to the multi-scale push-pull technique in [Grossman, et al. 2007]. This display system is adapted to benefit from high density clouds generated using super-resolution methods. Point clouds are rendered with push-pull interpolation in image space. A min-depth check and kernel density estimator are used to resolve edge discontinuities and remove occluded background measurements. The utilization of texture in-painting for both color interpolation and depth reconstruction provides the viewer with graphical continuity as well as proper occlusions, which standard point displays lack. In addition to fast rendering of huge datasets, the renderer allows the model to be updated in real time as new data arrives without costly re-meshing operations. The system can generate real-time (>30Hz) imagery at 1080p HD resolution on commodity (GeForce GTX 260) hardware with point clouds of greater than 5 million points.

6.1.4 Analysis

The results show that the method increases interpolation accuracy by up to 20% on the Bruceton Mine data, with an average improvement of 12%. The fisheye-spinner setup features density increases up to 70 fold, with an average of 40x increase in density (Table 9). Of note is that real resolution is created where LIDAR beam physics dictate a maximum angular resolution. This is apparent in 3D scanning mechanisms that actuate a planar sensor, where an increase in data collection time results in diminishing resolution returns. Lastly, a comparison of LDR and HDR imaging shows that increasing the dynamic range greatly improves the density and improves accuracy to a lesser extent. Due to the harsh artificial illumination in the naturally dark mine, saturation plays a significant role in the amount of usable data. Accuracy increases from HDR are the result of enhanced decimal resolution for surface normal estimation.

To validate that true information is being stored in the interpolated values, a sliding-window 15x15 pixel Pearson correlation was performed on the test data. As shown in Figure 62, the shaded image provides significant information about the ground truth that is not contained in interpolation. The fused range map correlates more than either source individually, concurring with the error estimation benchmark. While Diebel's method shows an infinitesimal increase in numerical accuracy (Table 8), it is not statistically significant. This is corroborated by almost equal amounts of strongly negative and positive correlation in the raw image data.

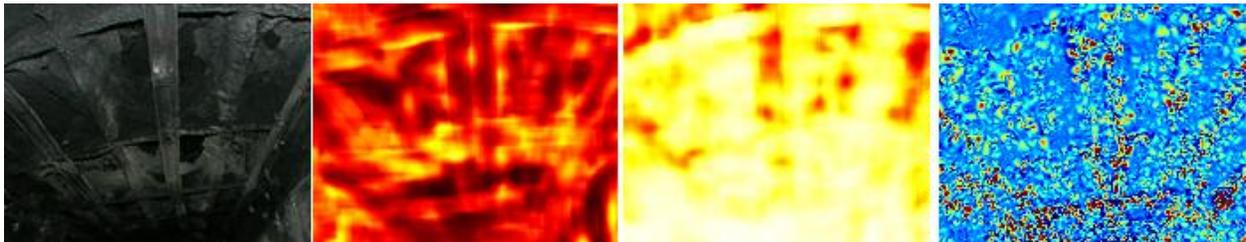


Figure 62. (Left to right) Roof supports covered in Shotcrete; Raw intensity to ground truth correlation; Shape-from-Shading estimates to ground truth correlation; and Reconstruction error reduction. Scale is brown to white over $[-1, 1]$ for correlation and navy blue (0.05m error reduction) to red (0.05m error increase). Discontinuity edges present the great amount of error increase.

The method encounters several drawbacks that prevent the fused result from achieving the same accuracy as LIDAR scans of equivalent density. Resulting range images are vulnerable to artifacts typical of raw interpolation, although to a lesser degree. Most reconstruction error occurs at occlusion edges where neighboring LIDAR points have large disparities. Regularization terms tend to over-smooth these edges and shading cues are ill-behaved due to cast shadows, among other reasons [Braquelaire, et al. 2005], [Worthington 2005]. Attempting to isolate these specific edges in the image is difficult due to image noise, lighting and material specific effects and is not addressed in this research (see [Torres-Mendez, et al. 2008], [Yang, et al. 2007]). Specularities in the environment were ultimately just a minor issue.

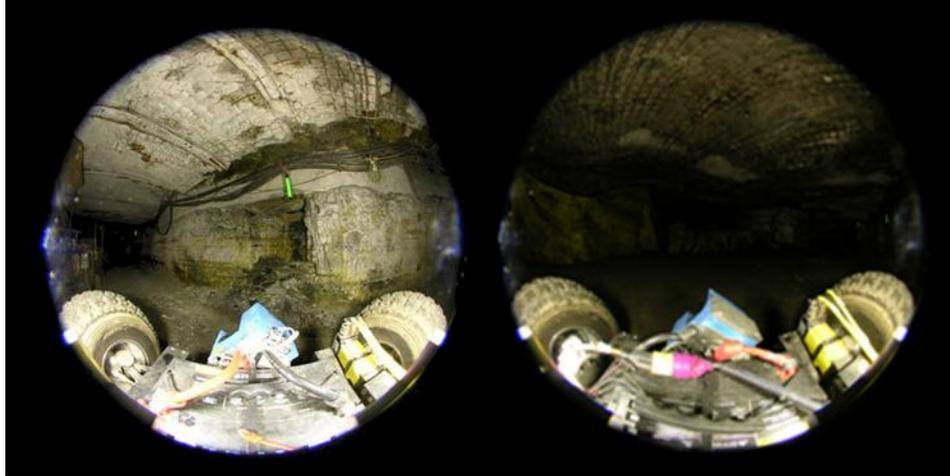


Figure 63. Scenes with Poor Results. Irradiance compensation breaks down when the robot approaches too near a wall (left). A large open area is not sufficiently illuminated by the robot's light source (right).

6.1.5 Discussion

A method was presented that fuses actively illuminated CCD imagery and LIDAR data. The method demonstrates increases in range accuracy of up to 20% on experimental data over interpolation and increases in measurement density of up to 70x using the experimental setup. The improvements are a result of calibrated imaging using additional knowledge of the image formulation model to reconstruct a 3D observation of the scene. This research demonstrated the efficacy of multi-sensor mapping systems as well as calibrated imaging for field robots.

Perhaps the greatest argument for range/image super-resolution is that it is easily bootstrapped to existing systems. Subterranean robots already require light sources for photography as well as range sensors for mapping and many high-throughput commercial scanners feature co-located cameras. The general use of illumination information for super-resolution is also applicable to other environments in the domain. Planetary robots are likely to encounter highly diffuse environments (i.e. Mars) or characterizable reflectances on bodies lacking scattering atmospheres (i.e. moon, asteroids). Such development is likely to also increase the safety of exploration and prospecting on the moon, where sensing is secondary to payload and comes at a premium cost.

6.2 Image-Directed Sampling for Geometric Modeling of Lunar Terrain

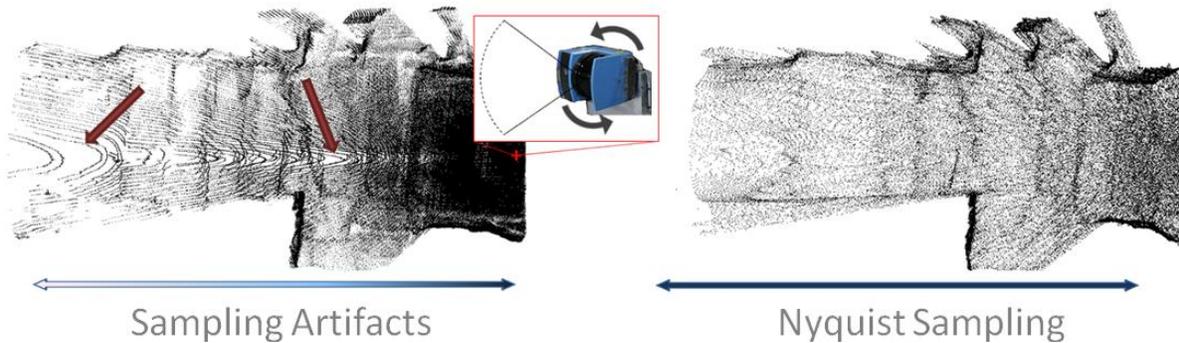


Figure 64. Approximate Spatial Uniformity from Adaptive Sampling.

Despite proliferation of motion-coupled industrial scanners in field robotics, there remain important applications which require actuated sensors and intentional sampling. These applications must consider the question of: "Where to sample the data?" Planetary exploration is perhaps the best example of the need to plan geometric sampling. Correct photometric and geometric classification of rocks and craters could enhance automated sample collection and manipulation for drilling on science missions. Steered and optically reconfigurable flash LIDAR is of great interest in automated planetary landing for its capability in handling a spectrum of ranges and scales. Lastly, cooperative orbital and ground mapping, where the robot itself is a bore-sighted, Dirac sampler, will likewise benefit from a saliency-based approach to resolving terrain obstacles first detected in aerial sensing [Jones, et al. 2012]. In these applications the importance of intelligent sampling is twofold: (1) the relative expense of range scanning places a premium on sampling well and (2) the requirements of model quality dictate the application.

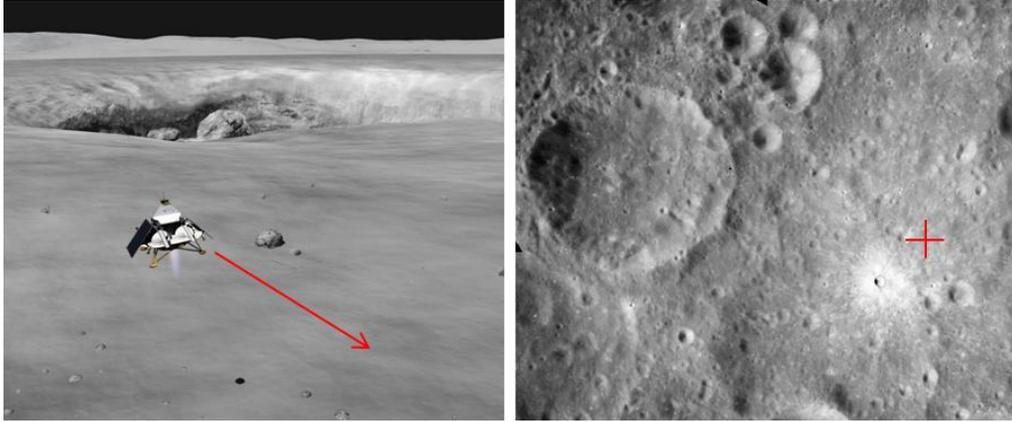


Figure 65. Automated Lunar landing is one possible application of image-directed sampling. A robotic lander (left) must model the near-field terrain to select the best landing site closest to a location of interest. Images generated during descent (right) can guide scanning with a laser altimeter to resolve regions of ambiguity. [Lander photo courtesy Red Whittaker, CMU; Apollo image AS15-M-0103 from NASA/JSC/ASU].

The quality of 3D reconstruction from point measurements is a function of the density, distribution and order of sampling. Only grid-based or uniform angular sampling strategies – which are commonly used – inform *a priori* the total number of readings required to scan a scene from a single viewpoint. In spite of this, these are poor sampling strategies for reconstruction. Triangulation of the resultant point clouds produces glaring artifacts like polygonal slivers and incorrect connectivity. Aliasing of grid-based approaches creates high gradient planes out of depth discontinuities. Wasteful and duplicate measurement increases sensing time for little information gain. These effects are particularly problematic in applications where the 3D detail is used in object recognition or understanding. One of the compelling motivations of this work stems from the observation that sensors which are not limited by grid sampling are the same (low-throughput) sensors which would most benefit from sampling in a principled manner.

Cameras are natural complements to range sensors; with the principles of Lumenhancement, it is possible to utilize color imagery to *direct* range sensing of a scene to avoid artifacts. At the core of this Image-Direct Sampling (IDS) approach is the observation that under planetary appearance assumptions, the frequency content of color images correlates strongly with that of geometry images. By sampling the scene with Nyquist-sufficient density distribution and by utilizing camera information to plan such a sampling, vast improvements can be made in both the quality and efficiency of range sensing.

This case study identifies frequency features for use with the image-directed scanning approach, compares feature performance based on reconstruction metrics and demonstrates new sensor designs and visualization techniques which utilize this paradigm. The advantages of diffuse planetary environments are discussed in tailoring domain-specific salient image features. Lastly, the aforementioned applications are used to test the efficacy of the approach.

6.2.1 Foundational Work

The image-directed sampling approach can be seen as a complementary approach to that detailed in the super-resolution study and in [Wong, et al. 2009]. Specific motivations for sampling arise from the observation that the quality of geometric fusion depends as strongly on the quality of the underlying range data as it does on the image-based shape estimation. By affecting the nature of range samples, this approach can be utilized in tandem with a spectrum of traditional fusion techniques [Diebel, et al. 2005] to significantly enhance range models.

Research towards determining and generating optimal point samples is well documented. Work of particular note includes [Martinez, et al. 2007], which devised a method to reduce the post-processing time and increase the convergence probability of multi-scan alignment by storing only high-saliency points from scans. Related multi-view extensions of the sampling problem include the body of next-best-view work [Shahid, et al. 2007], [Mark, et al. 2010]. Frequency content is also distinguished as a strong predictor of novelty in exploratory map building in [Thompson, et al. 2008], specifically for planetary terrains.

Wavelets and other multi-scale frequency features are widely used in data understanding and reconstruction. They include approaches to LIDAR [Wei, et al. 2006] and intensity images [Loupias, et al. 2000], [Achanta, et al. 2009]. In particular image based approaches have been successfully used in interest seeking for planetary science [Dunlop, et al. 2007]. Range and intensity fusion with wavelets has also been studied for forest region mapping [Wang, et al. 2007].

While image-directed sampling with Lumenenhancement draws inspiration from prior work, it distinguishes itself in two important aspects. Prior approaches have enhanced unstructured range data using image content in a post-processing manner, this work tightly couples image analysis in the range *acquisition* phase. The scanning process is redesigned from the ground up to produce models which exhibit density characteristics - such as spatial uniformity - that are intrinsically advantageous for fusion or volumetric reconstruction from a single view. Thus, this approach supplements many of the multi-

view techniques described above. Secondly, this work is motivated by analysis of surface reflectance in planetary appearance which can constrain image features to physics-based analogues. The properties of features are not explicitly considered for detection, recognition or other high-level purposes here, as these are well-studied.

6.2.2 Frequency-Based Modeling

Frequency is one of the most effective predictors of information content in a signal. Continuous time-varying signals can be decomposed into their frequency components by utilizing a number of different transforms to the frequency domain. There is no exact definition of this domain, but transforms generally have desirable properties that indicate the change-over-time of the signal; other closely related concepts include “saliency” and “compressibility”. One such common procedure is Fourier analysis, which decomposes signals as a (possibly infinite) sum of sinusoids of varying periodicity. Lower frequencies are data content that represent bias offsets and trends while higher frequencies contain details and sensor noise. The envelope of frequencies describes how the signal changes at all scales.

The frequency components of range models cannot be directly measured. The only sensors that exist for range modeling are Dirac samplers such as Time-of-Flight and Triangulation, both which produce point estimates. It is therefore prudent to consider the Nyquist sampling theorem when reconstructing a scene from sub-sampled points. Given a band limited signal of bandwidth B defined as its Fourier Transform:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft} dt \quad (6.14)$$

$$X(f) = 0 \quad \forall |f| > B$$

the original signal can be reconstructed exactly by a sub-sampling with a uniform frequency of rate f_s and interval T , where:

$$f_s > 2B \quad (6.15)$$

$$T = \frac{1}{f_s} \quad (6.16)$$

Paraphrased, the theorem states that a sufficient condition to prevent aliasing is a uniform sampling rate at twice the highest frequency present in the signal [Forsyth, et al. 2002]. In signals practice, a fixed

sampling rate much greater than the Nyquist frequency is often used, but that approach is often not possible in LIDAR modeling for the range of resolutions or the fixed number of samples desired.

It is common knowledge that the areal *density* of points affects the reconstruction accuracy; an ideal plane is perfectly described with 3 non-collinear points, while a step-edge cliff requires many points. The Nyquist theorem applies more generally to non-uniform samplings to say that a signal can be perfectly reconstructed if the average sampling frequency satisfies Nyquist [Landau 1967]. Moreover, perfect reconstruction is possible from severely sub-Nyquist spatial samples, if the signal has *known* compact support and shift invariance in the frequency domain [Aldroubi, et al. 2001]. However, this sampling has increased error sensitivity to a uniform approach [Venkataramani, et al. 2001].

It is impossible in practice to know if a scene satisfies these conditions a priori. Furthermore, noise and sensor discretization prevent exact frequency content from ever being recovered with sampling. Thus, only effective use of samples can be made by apportioning low and high information areas with some sample weight distribution and relative regard for Nyquist. This corresponds to oft-opposing objectives for modeling in a frequency sense: (1) accuracy of volumetric reconstruction and (2) localization of high spatial resolution detail to important areas. A range of sampling weight distributions should therefore be explored for application-specific reasons, rather than attempting a theoretically optimal approach.

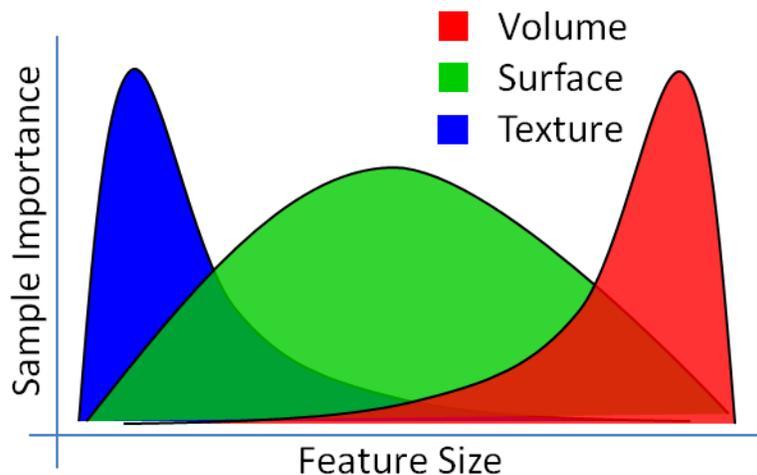


Figure 66. Example Spatial Sampling Frequency Distributions based on application intent.

There are several approaches for selecting the spatial distribution of sampling. The advantage of spatially uniform sampling is that the distribution of data bounds the maximum volume error in a polygonalized model to a function of the sample density. As true structure between two known points is

unlikely to differ significantly from a linear gradient, spatial uniformity imposes a simple inverse-scalar relationship between the number of points and the error. The restriction of a regular surface samples also reduces the probability of triangulation artifacts in meshing. Another distribution of interest is frequency-weighted, which minimizes the frequency spectrum error of reconstruction. Existing sampling methods do not consider the *importance* of regions in a frequency sense; salient feature detection is enhanced by contrasted but invariant discriminativity over the background. For applications like tracking or identification, it is more desirable to sample surfaces appropriate to their detail and independent of position or range. However, increased sampling of high frequency regions can also magnify noise.

The sampling density distribution can be tuned to balance the tradeoff between global model consistency and local feature preservation to suit the specific application. This succinctly describes the approach to frequency-based modeling advocated here. Ultimately, both distributions converge at the macro scale, where the amplitude of surfaces is comparable to the sensing range.

6.2.2.1 Wavelets and Range Images

Two dimensional range images (and their geometry image cousins [Gu, et al. 2002]) are common manifold representations of 3D models. The wavelet transform is a simple and elegant representation of the spatially varying frequency content in such grid-based data. By successive subsampling and decomposition in a cascading filter bank, an image pyramid of *detail* (high frequency) and *approximation* (low frequency) coefficients is produced. If the filter satisfies specific requirements, the levels of the pyramid are approximations to the frequency content [Daubechies 1992]. The total energy, the sum-squared “detail” coefficients is one indicator of the information distribution [Rosso, et al. 2006], [Achtuthan, et al. 2010].

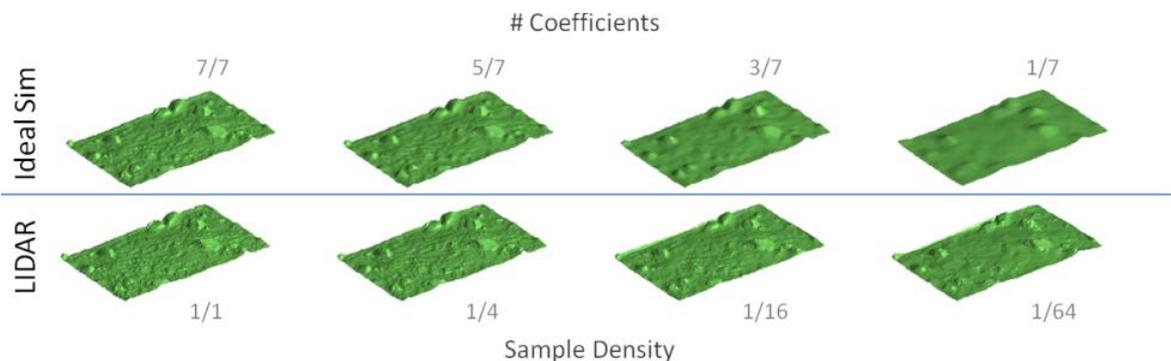


Figure 67. The frequency content of a region is proportional to local sampling density. This relationship is illustrated empirically by decomposition of a range image with wavelets (top), and reconstruction by downsampling a point cloud (bottom).

In the decomposition of range images, the lowest levels of a wavelet pyramid can be seen as the mountains and valleys of terrain, for example, while successively higher levels are increasingly representative of local features, like much smaller rocks. The "approximation" component provides the most of the shape and reconstruction volume in a scene. The Wavelet Transform will be the vehicle for frequency analysis in this approach.

6.2.2.2 An Image-Directed Approach to Sampling

Frequency-based modeling requires that the spatially localized information content of the scene be known *a priori* in order to plan a variable sampling density [Venkataramani, et al. 2001]. This approach is self-defeating as knowledge of the underlying frequencies requires scanning the scene modulo the maximum working resolution. A prior over the Nyquist frequency distribution is sought, which can guide spatially varying sample acquisition. Such a prior can be estimated with a complementary sensor, such as a camera, which can instantaneously capture information with the same perspective as a range image.

Pixel intensity is a composition of three-dimensional geometric, material and illumination properties projected onto a two-dimensional plane [Hom 1986]. Estimation of shape from images is woefully underconstrained in the general case; however, a Lumenhancement approach can be taken in planetary environments with constrained appearance. This work considers intensity images, whether raw or feature-transformed, which correlate significantly with scene geometry. Features from these images are viewed as noisy approximations of high resolution range or geometry cues. The content of these images holds minutia about the spatial and frequency properties of the high density geometric sampling that cannot be surmised from a low density sampling alone.

Recall the rendering equation in section 2.3, equation (2.1):

$$L_o(x, w_o) = L_e(x, w_o) + \int_{\Omega} f(x, w_o, w_i) L_i(x, w_i) (w_i \cdot \hat{n}) dw_i \quad (6.17)$$

As shown previously, it can be reduced to manageable form (6.22) using the assumptions of planetary environments.

$$E(x) = L_o(x, w_o) = L_e(x, w_o) + \sum_k f(x, w_{o,k}, w_{i,k}) L_{i,k}(x, w_{i,k}) (w_{i,k} \cdot \hat{n}) \quad (6.18)$$

$$E(x) = \sum_k f(x, w_{o,k}, w_{i,k}) L_{i,k}(x, w_{i,k}) (w_{i,k} \cdot \hat{n}) \quad (6.19)$$

$$E(x) = \sum_k \frac{\rho}{\pi} L_{i,k}(x, w_{i,k}) (w_{i,k} \cdot \hat{n}) \quad (6.20)$$

$$E(x) = \sum_k \frac{\rho}{\pi} I_k (w_{i,k} \cdot \hat{n}) \quad (6.21)$$

$$E(x) \propto (w_{i,k} \cdot \nabla x) \quad (6.22)$$

These assumptions are *barren* (6.21), *dry* (6.20), *rocky* (6.19) and *simple illumination* (6.18). Thus, the image content is related to the geometry by a differential operation (6.22). The frequency relationship between a signal parameterized as $z = f(x, y)$ and its gradients $p = f_x$ and $q = f_y$ is given by [Frankot, et al. 1988], [Agarwal, et al. 2006]:

$$Z_f(u, v) = -i \frac{uP(u, v) + vQ(u, v)}{u^2 + v^2} \quad (6.23)$$

where $Z_f = F[f](u, v)$ denotes the Fourier transform of the geometric height field, and (P, Q) the Fourier transforms of its gradients. Therefore, it is possible to recover a Nyquist-optimal sampling of scene geometry given prior information from a perfect intensity image of the scene. However, the differentiation operation does have the undesirable effect of magnifying noise, which is often a high-frequency component. For this and reasons mentioned previously, heuristic approaches for artifact reduction are advocated. Moreover, this analysis assumes minimal cast shadowing (similar to the barren constraint), as these introduce high-frequency edges into the image.

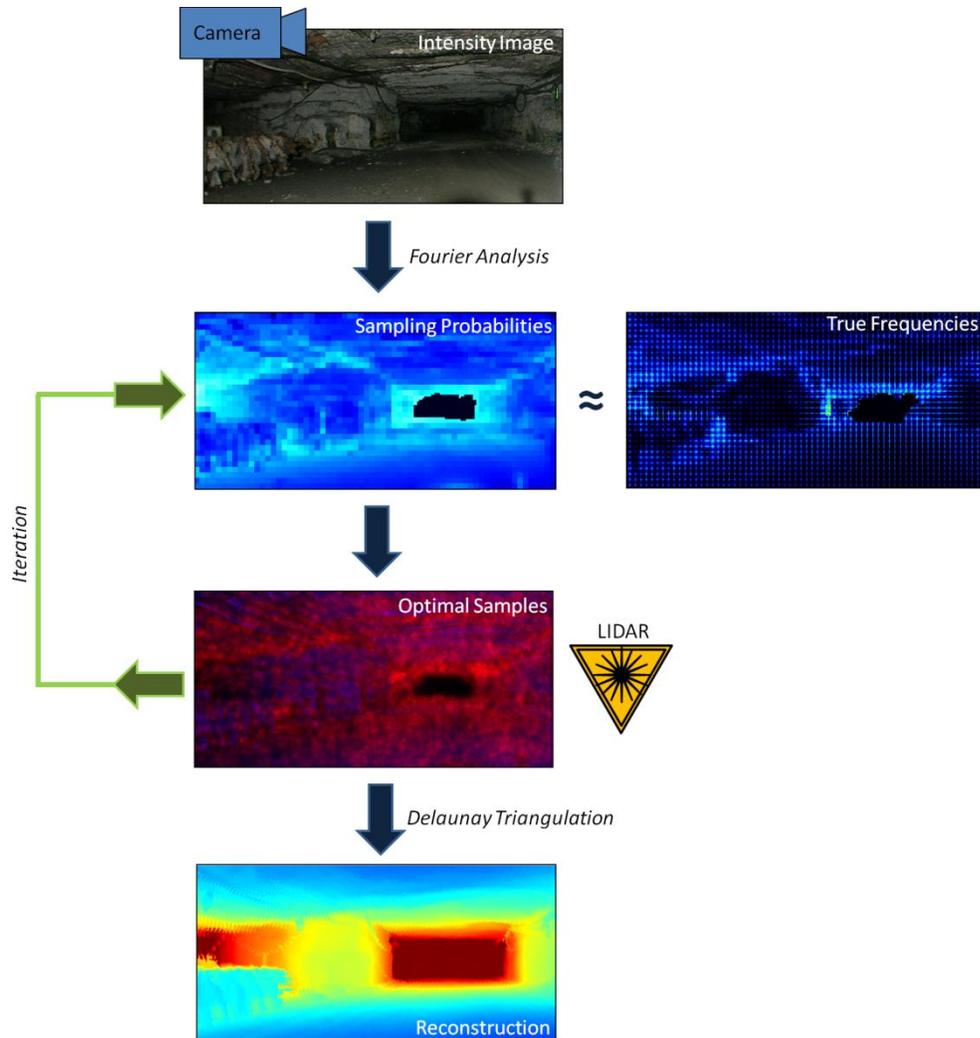


Figure 68. Image-Directed Sampling Process. (1) Intensity Image features generate a sampling probability map. This map approximates the true surface frequencies of the scene. (2) Optimal samples are selected using the probability map and acquired with a manually guided LIDAR scanner. (3) The process is iterated until stopping conditions are satisfied. (4) Triangulation of the raw LIDAR points produces a 3D reconstruction.

Figure 68 provides a high level illustration of the approach taken here. The optimal reconstruction problem presented here is viewed as choosing samples $x_k = \langle s_1, \dots, s_k \rangle$ from all possible samples $s_i \in X$, such that the reconstruction error is minimized: $e = \arg \min_s \|t(x_k) - X\|$ under some interpolating function t . The objective is to determine a suitable prior using features from intensity images (i) and learn a mapping such that $f(i, X) \rightarrow \hat{x}_k$, where $\hat{x}_k \sim x_k$ is minimized. The thesis research coins the term *image-directed sampling* for this ensemble framework.

6.2.3 Modeling for Planetary Robots

Validation of the image-directed sampling approach is conducted in the specific context of lunar robotics. In addition to optical domain relevance for Lumenhancement, applications provide several opportunities to explore sampling. Private enterprises, such as the Google Lunar X-prize, have renewed interest in automation of landing and exploration on the moon. There is particular emphasis on creating high quality maps and models of the moon using robots as precursors to humans. The approach is explored in both aerial sensing and ground-robot sensing on the moon to demonstrate the robustness of the technique to scale.

6.2.3.1 Lunar Aerial Mapping

This section describes validation of the approach on simulated, aerial-scale lunar terrain²⁹. This terrain is procedurally generated in Blender™, utilizing randomized fractal crater and rock distributions taken from the Surveyor 6 mission [Heiken, et al. 1991]. The data includes independent overhead depth and RGB values for each voxel in the scene at 5vox/m. Three different scenes were generated and are shown in Figure 69.

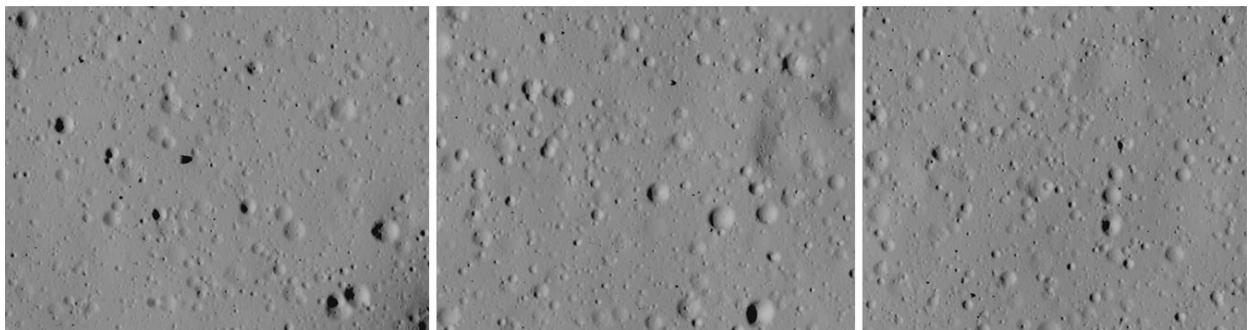


Figure 69. Image features in diffuse environments, such as the moon, correlate strongly with geometric features. These three simulated lunar terrains are utilized in this work.

The datasets simulate lunar sensing during terminal descent, at a scale 50m above the ground and were generated for related research. An automated lander, for example, might identify geometric hazards (rocks and craters) and divert to a suitable landing spot. Image-feature based methods have been proposed to identify these hazards, including detection of shadows [Hata, et al. 2004]. However, in this scenario a gimbaled altimeter or boresight flash LIDAR can also work in tandem with image-based methods to acquire a full 3D model.

²⁹ Simulated lunar terrain is joint work of the author and documented in [Jones, et al. 2012]

Simulation of the sampling process includes selecting samples (voxels) from the highest resolution depth map, triangulating the subsamples and linearly interpolating to create a depth map at the native resolution. This depthmap is then compared against the ground truth digital elevation map to produce reconstruction error scores. Artifacts such as sensor noise are not considered in this analysis. Several scanning strategies were compared on this dataset. They include:

Uniform Grid – An $N \times M = k$ element uniform grid sampling across the scene, rounded to voxel edges.

$$w = \mathfrak{I}(\mathbf{I}) \quad (6.24)$$

where \mathfrak{I} is the dirac comb function for grid locations.

Uniform Random – Random sampling of k elements in the scene with each voxel receiving uniform weight.

$$w = c \quad (6.25)$$

Gradient Weighted – Random sampling weighted by the 2-norm of the partial image derivatives, a precursor to image-based edge detection. Calculated using the matlab command `gradient`.

$$w = \left\| \nabla_x I + \nabla_y I \right\|_2 \quad (6.26)$$

where $\nabla_x I$ and $\nabla_y I$ are the image partial derivatives in the x and y direction respectively.

Wavelet Weighted – Random sampling weighted by the sum-square of the wavelet detail coefficients, normalized by subregion size. The wavelet transform is generated with DB2 wavelet using the command `wavedec`.

$$w = \sum_{j=1}^N \frac{1}{2^j} (D_j)^2 \quad (6.27)$$

where j is the pyramid level and D_j are the detail coefficients at level j .

Entropy Weighted – Random sampling weighted entropy in a 9x9 image neighborhood centered about the query pixel, calculated with the matlab command `entropyfilt`. This transform is frequently used as a texture cue.

$$w = E(I) \quad (6.28)$$

Three separate metrics were used to compare the reconstruction error between sampled depth maps and ground truth. These metrics each emphasize a different requirement of the model.

Mean Squared Reconstruction Error (MSE) – Penalizes for any large difference between the reconstruction and the ground truth.

$$E_{MSE} = \frac{1}{N \times M} \sum_{i \neq k} (t(\hat{x}_k) - x)^2 \quad (6.29)$$

where $t(\hat{x}_k)$ is an interpolation of k samples from x .

Gradient Weighted Reconstruction Error (GRAD) – Penalizes for incorrect reconstruction at depth discontinuities.

$$E_{MSE} = \frac{1}{N \times M} \sum_{i \neq k} \lambda(\nabla x) \|t(\hat{x}_k) - x\|_1 \quad (6.30)$$

where λ is a weighting function on the range gradient.

Saliency Weighted (SAL) – Assigns higher weight to in reconstruction salient features such as rocks and craters correctly. The saliency map is generated using using the Saliency Toolbox [Walther, et al. 2006].

$$E_{SAL} = \frac{1}{N \times M} \sum_{i \neq k} \lambda S(x) \|t(\hat{x}_k) - x\|_1 \quad (6.31)$$

Random samples were generated 20 times using each strategy and the resultant scores were averaged. In the case of the grid strategy, the grid was shifted horizontally and vertically so that sampling did not always occur in the same voxels or on the edge of the scene. Optimal linear coefficients were found using `fminsearch`. The experiments were performed with fractional sampling density of the ground truth resolution in log spaced increments, i.e. $k = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots, \frac{1}{2048}$.

Table 11. Performance of Sampling Strategies on Simulated Aerial Terrain

Scene #1	Random	Grid	Gradient	Entropy	Wavelet
MSE	1	1.02	0.88	0.87	1.11
GRAD	1	1.1	1.10	1.13	1.12
SAL	1	0.93	0.93	0.88	1.07
Scene #2	Random	Grid	Gradient	Entropy	Wavelet
MSE	1	1.01	0.87	0.85	0.99
GRAD	1	1.05	1.12	1.16	1.12
SAL	1	0.82	0.96	0.95	1.05
Scene #3	Random	Grid	Gradient	Entropy	Wavelet
MSE	1	0.91	0.83	0.80	1.21
GRAD	1	0.98	1.02	1.06	1.01
SAL	1	1.02	0.82	0.79	1.31

The results of the sampling simulation are detailed in Table 11. The scores are given as the mean ratios of the error between random sampling and the strategy in question, weighted by inverse sample density (6.32) where T is the total number of voxels to select from and k is the number of samples selected.

$$score = \frac{1}{\sum k} \sum_{k \in S} \frac{T}{k} \cdot \frac{e_{rand|k}}{e_{strategy|k}} \quad (6.32)$$

Thus, scores represent an improvement multiplier over random sampling. Reweighting by inverse density accounts for performance across the entire logarithmic sampling densities of interest while discounting minute reconstruction noise due to lack of selection replacement at the highest density trials.

It is noted that in very sparse samplings ($k < 15000$), wavelets performed the best by far in all metrics (see sample curve in Figure 71). This is particularly promising when high reductions in the amount of geometry acquired are necessary. Wavelets perform more modestly (20% improvement) given higher numbers of points, as low frequency regions seem to be undersampled. Gradient and entropy approaches both perform relatively poorly. Gradient weighting fails because the sharpest edges occur in shadowed craters, not depth discontinuities. A similar problem occurs with entropy, which utilized a fixed scale and neighborhood. While the wavelet transform is also susceptible to overweighing shadow features, the multi-scale capability also captures the crater and rock features. There does not appear to be any significant difference between random and grid based sampling approaches (the grid approach

comes out marginally ahead). This is perhaps due to the fact that while grid sampling is subject to bias and aliasing, this is balanced by the possibility for large holes in purely random sampling. A better pseudo-random sampling approach for all these features might first compute a Delaunay triangulation within similarly valued regions to ensure well-behaved subsamples.

6.2.3.2 Lunar Ground Mapping

Unlike aerial sensing, which is difficult to scale for lab experimentation, high-fidelity robot scale terrain can be constructed with readily available materials. This section describes experimental verification of the approach on the lunar-like moonyard terrain (see section 5.2).



Figure 70. Lunar Analog Terrain with simulated overhead sun light

The terrain was placed in a light simulator frame with calibrated sensor and sun-scale light source positions for data collection (an overhead sun position was used). Range data was collected with a survey-grade phase-shift LIDAR by scanning from multiple positions to ensure adequate resolution and elimination of range shadows. Likewise, HDR imagery was collected from multiple locations utilizing DSLRs. The raw LIDAR point cloud, consisting of 5million points, was colored with RGB imagery, cropped to the edge of the sandbox and averaged into a 971x1674 voxel (1.6M samples) orthographic digital elevation map. Figure 76 in the next section shows the data acquisition setup.

As the scanner does not have independent sampling control, it was necessary to scan at maximum resolution, voxelize into a uniform grid and sub-select samples post-measurement for testing.

Quantization into lower resolution voxels enables independent readings for each position without interpolation and provides resilience against the natural angular biases of the sensor.

Table 12. Performance of Sampling Strategies on Lunar Terrain

	Grid	Random	Gradient	Entropy	Wavelet
MSE	1.1	1	1.09	1.15	1.33
GRAD	0.82	1	1.21	1.32	1.42
SAL	0.96	1	1.2	1.30	1.50

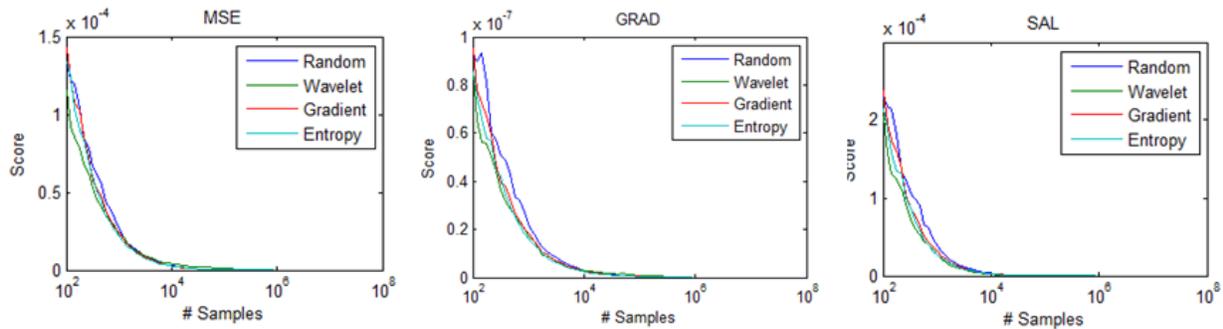


Figure 71. Comparison of selected strategies over all subsample amounts and metrics on lunar terrain model. Visually, wavelets are the best performers (green).

The sampling strategies described in the previous section were tested and the same metrics described are used in this analysis without change. The results of the experiment in artificial lunar terrain are summarized in Table 12. The wavelet-based sampling is clearly the best performer in this experiment (~40% improvement). Curiously, all of the strategies performed significantly better than in the simulated aerial imagery. It is believed that this effect is due to the strong cast shadows in the simulated imagery, while this experiment (though using real data) featured softer shadows and few large, negative features. While results are promising, it must be noted that this single dataset cannot be representative of the entire spectrum of lunar appearance and geometry and additional testing is required.

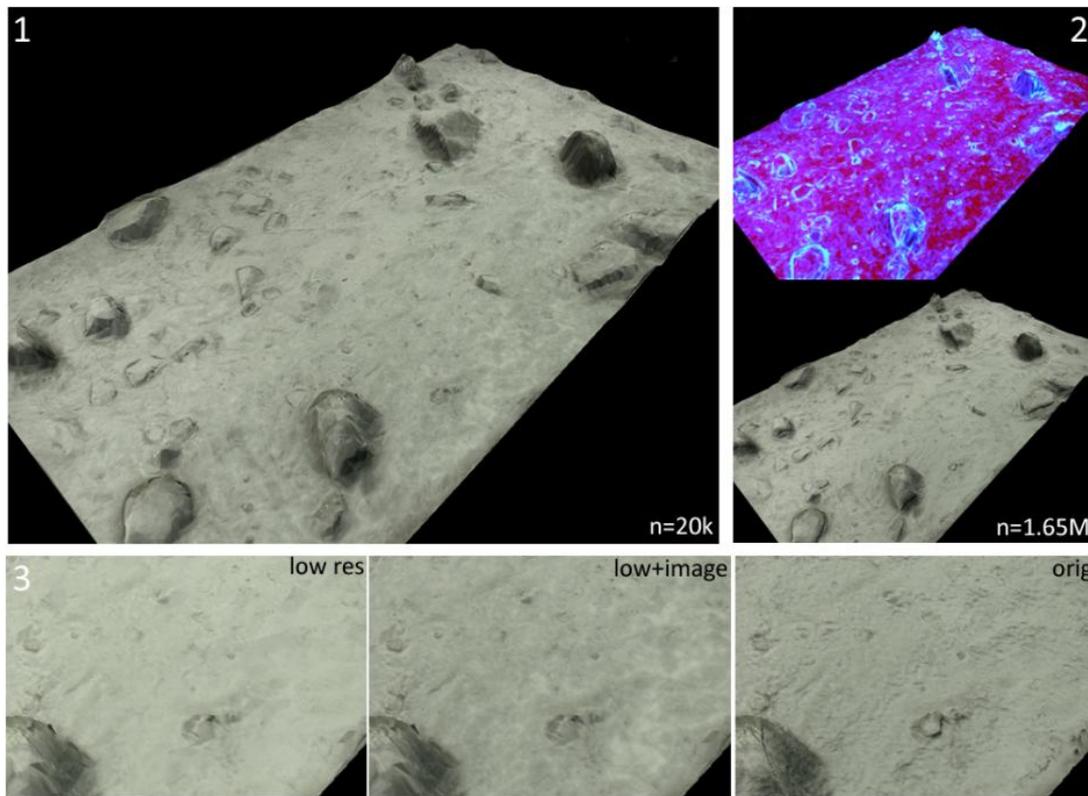


Figure 72. Lunar terrain model generated with image-directed sampling. (1) An extremely sparse sampling rendered with roughness-modulated Oren-Nayar BRDF retains useful geometry and is visually appealing. (2) The frequency map of the terrain generated using wavelet decomposition. (3) Detail of roughness visualization demonstrates that a data deficient model can retain much of the optical appearance of the highest resolution model.

Visual results of terrain reconstruction using wavelet-weighted, image-directed samples are illustrated in Figure 72. Figure 72-1 shows a rendering of the terrain with 20 thousand sample points, representing an 80x density reduction from the ground truth model of 1.6 million points (Figure 72-2, bottom). As a result of well-placed samples, the macroscopic geometry is surprisingly well-formed for such a sparse mesh; there are few glaring differences when compared to the ground truth. The downsampled model is rendered with an inferred roughness from the wavelet decomposition. High frequency features estimated in the image are "painted" onto the sparse geometry by modulating the roughness term in an Oren-Nayar vertex shader proportional to the energy [Oren, et al. 1994]. A threshold can be applied to ensure major geometric features remain metrically true. This approach to visualization, coined *adaptive surface frequency*, conveys the high resolution surface characteristics to the viewer while preserving mesh compactness with minimal computation. Figure 72-3 further illustrates this technique. With roughness inference turned off, the surface is an aliased collection of triangles, but rendering with roughness infuses many of the characteristics of the true surface.

6.2.4 A Class of Image Directed Scanners

Prior sections demonstrated the method on simulated data or utilized selective sampling on pre-captured high resolution data. This section discusses two hardware implementations of the approach: one on traditional actuated sensing and the other on a novel type of independent sampling sensor.

6.2.4.1 Sampling with Constrained Motion

Thus far, a discrete grid based approach has been compared to random sampling approaches. Significant improvement was shown when random sampling is combined with image frequency estimates. However, it can be argued that this comparison is unfair. Few existing range sensors can acquire random, independent samples. Most of the applications that are considered in this work would use steered beam sensors if taken at the current state of sensor art. *Ferret*, a borehole-deployed underground inspection robot (Figure 73), which utilizes programmable steering is explored here. This robot servos a single beam LIDAR much like a lunar lander might steer a laser altimeter to scan the terrain, thus the two utilizations (and their appearance spaces) are very similar despite differing applications.

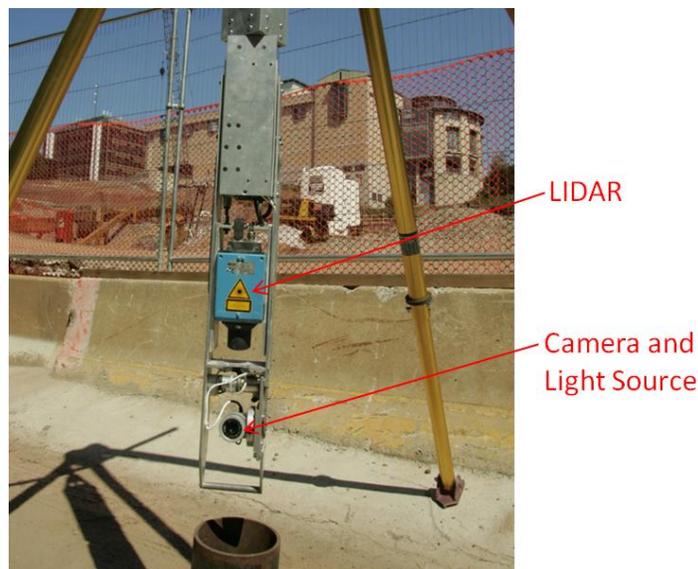


Figure 73. A borehole-deployed underground inspection robot scans with a single beam LIDAR by actuating pan and tilt axes. Mechanical steering constrains the placement of the next sample based on the position of the prior sample and limits on rotational velocity. An independently control camera can image the scene with an LED light source.

There are three classes of actuation for single-viewpoint scanning beam sensors. From a single viewpoint, two motion axes (corresponding to spherical coordinates) are required to sample the

hemisphere. Each of these axes can rotate in a single direction from one limit to another or oscillate between limits in the field of view. Unidirectional motion on both axes produces a three dimensional “coil” of readings and is energetically efficient, but mechanically complex. Oscillating motion of either axis produces a raster-type scan. Archetypical examples include tilting or “nodding” scanners. Lastly, oscillation of both axes produces a spiral-type scan.

Independent sampling (actuation of the beam to randomly selected measurement locations) is also possible in some setups. This strategy is extremely inefficient for mechanical scanners as it does not exploit synchronized parallelism between actuation and the measurement clock. Figure 74 illustrates possible trajectories on an example depth map.

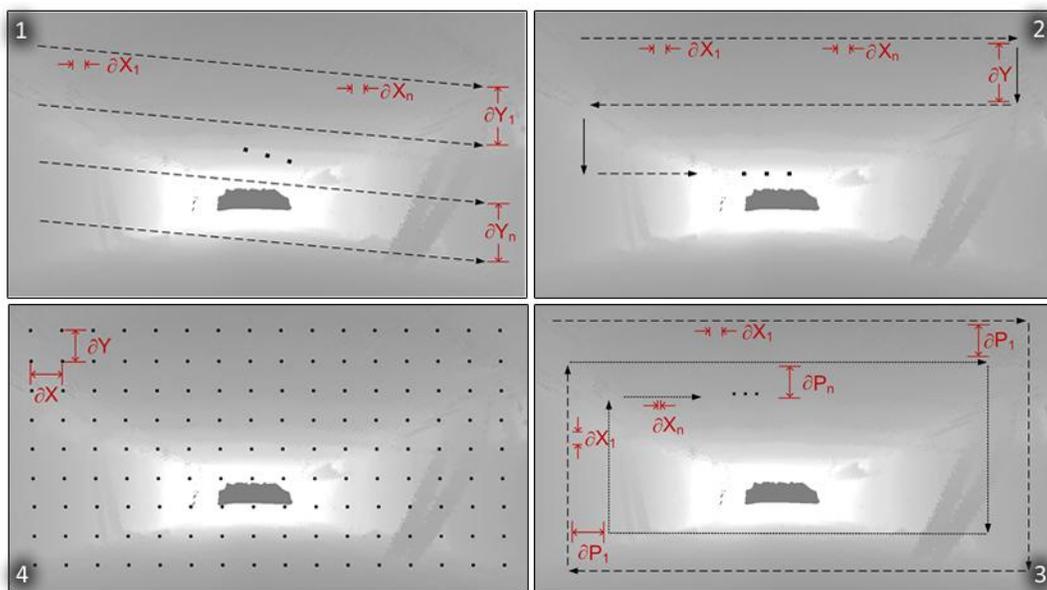


Figure 74. Actuation Trajectories simulated on range images of a scene. (1) Linear scanning produces a coil of data parallel to the scene. (2) Raster scanning oscillates the sensor head on a single axis while maintaining motion on the orthogonal axis. (3) Adaptive spiral scanning oscillates both axes, varying the tightness of the inner spiral based on data collected along the outer spiral. (4) Independent sampling selects locations at random for acquisition. This may require erratic actuation or an independent sensing array.

These approaches can be enhanced with data-dependent adaptive sampling strategies where previously measured data informs the placement of future data. Such strategies could be used to approximate collection of spatially uniform samples on the scene or target a variety of reconstruction properties. For example, by simply increasing the velocity of the tilt motion of a 1-DOF sensor in the near-field and decreasing velocity in the far field, an adaptive raster strategy can generate a more desirable model of the mine scene than data-independent approaches (Figure 75). However, in this case, sampling bias is

still present on the horizontal axis and a tradeoff exists between bounded acquisition time and uniformity of density in unknown scenes [Omohundro 2008].

However, iterative refinement is utilized here to apportion the scene into manageable regions that are each scanned in a naïve manner. Each region has a single uniform raster density and the scanner head stops collecting data briefly while moving between contiguous regions. Multiple passes can be utilized to selectively refine regions of high importance. Models created in this manner are a quantized approximation of random sampling.



Figure 75. Iterative refinement can be utilized in motion-constrained scanning to enable variable density sampling in regions while respecting the grid nature of actuation (1). A mine scene is scanned with Image-Directed Sampling using the *Ferret* borehole robot (2). Uniform surface sampling of the mine scene (bottom) illustrates much less aliasing than traditional naïve scanning.

In addition to frequency analysis, artificial illumination can be utilized underground for Lumenenhancement. Assumption of tunnel nature and source fall-off can be utilized to estimate the macroscopic shape and orientation of the environment to plan uniform-surface sampling. Surface normal distributions from SFS reconstructions along with material segmentation can cross-check

locations of high geometric frequency. Occlusion detection from multi-flash imagery can precisely localize sampling of edges. Figure 75 demonstrates scanning of an underground scene using iterative refinement with source fall-off and wavelet decomposition as sampling weight priors.

6.2.4.2 Independent Sampling with Structured Light

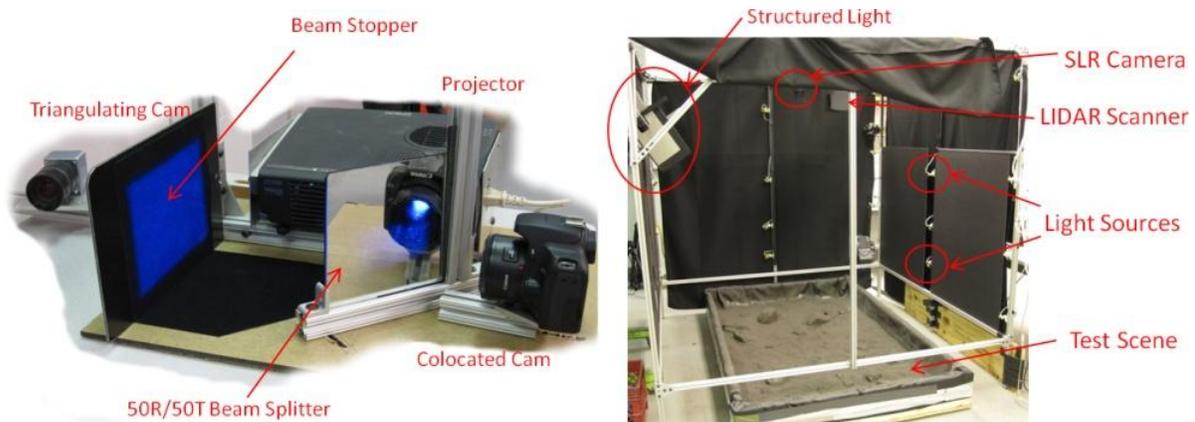


Figure 76. An image-directed scanner consisting of a high resolution camera optically co-located with a DLP projector (left). Experimental setup utilized in this paper showing mounting positions of all the sensors (right).

It is of particular interest to consider the class of sensors that provide pixel-independent sampling, as they are capable of implementing the approach in the purest form. Perhaps the most obvious and common of these are structured light sensors based on digital projectors. Though, groundbreaking work has been conducted in camera-integrated smart laser scanners [Cassinelli, et al. 2010].

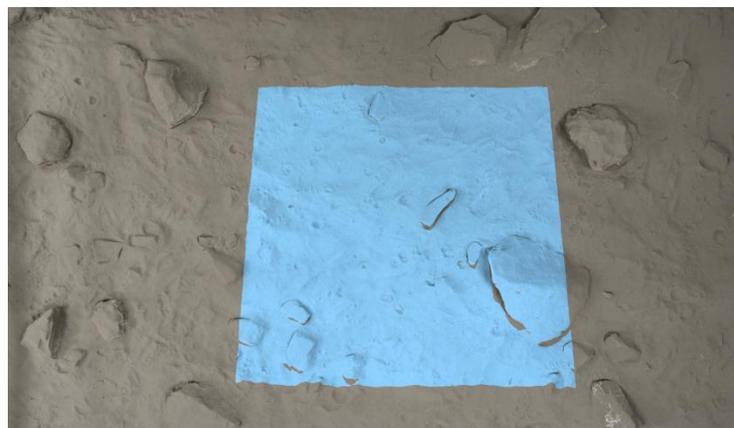


Figure 77. Structured Light scan volume on lunar test terrain.

A structured light scanner was modified to develop a proof-of-concept image-directed scanner. Figure 76 illustrates the optical configuration of this device. A high resolution DSLR camera is optically co-located with the projector through the use of a half-silvered mirror. The system is calibrated such that each pixel in the camera corresponds to an exact outgoing ray of projected light. This configuration enables the system to sample the exact location viewed by the camera without shadowing. Finally, a second camera is optically offset from the other two devices and measures depth via triangulation. This setup is closely related to that presented in [Fujii, et al. 2009], which is a co-located system for augmented reality. Image-directed sampling is implemented on this scanner as follows: the projector first lights all pixels so that the scene can be imaged by the DSLR and analyzed; afterwards, scanning proceeds in either of the two modes.

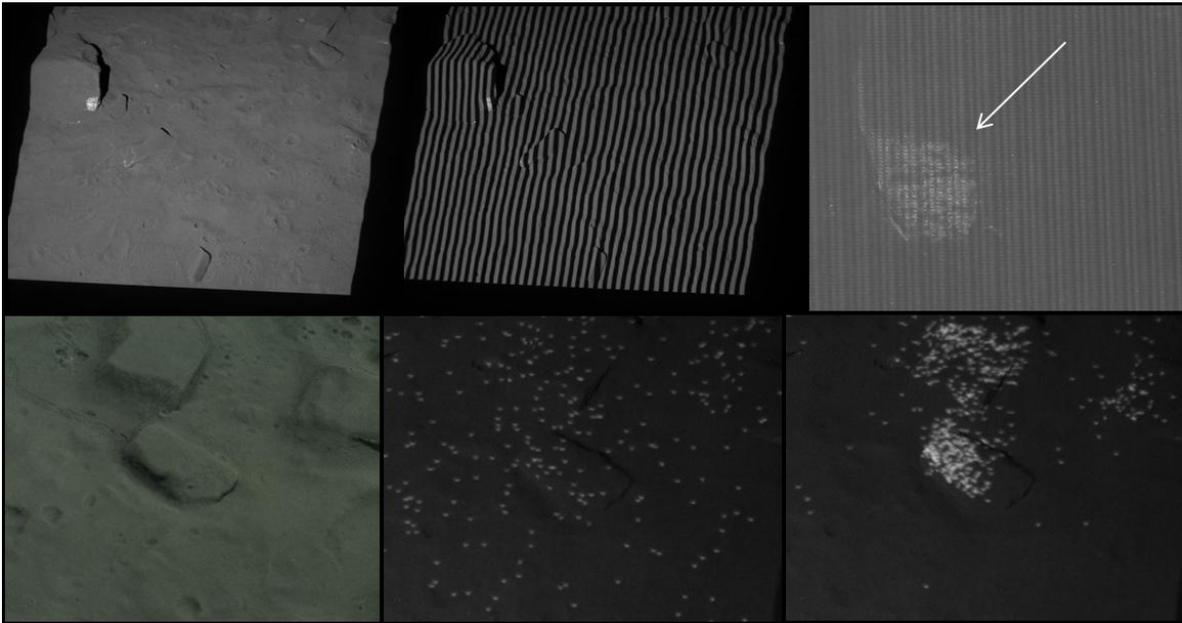


Figure 78. Linear gray code scanning of the lunar terrain causes errors at the finest resolution due to differences in underlying scene albedos (top). Image cues can direct the scanner to resample salient regions such as rocks a pixel at a time (bottom). Samples are shown simultaneously illuminated for illustrative purposes; scanning occurs one dot at a time.

The first mode is *ambiguity enhancement*. In temporally-coded stripe scanners, the highest resolution stripes often result in ambiguities or errors (illustrated in Figure 78 on the lunar test terrain). Among the sources for these errors include the albedos and complex reflectance of materials present in the scene to the resolving power of the camera lens. Gray codes can alleviate this problem, but result in lower effective resolution when such errors occur [Salvi, et al. 2004]. Image-directed sampling can be applied to identify regions which require dense, accurate samples, which can then be resampled one pixel at a

time (reducing errors). The bottom row of Figure 78 demonstrates salient features (rocks) that require more samples. The locations of these samples are all illuminated by the projector for illustrative purposes, but are actually scanned individually. This method enabled a 25% improvement in reconstruction error on the lunar terrain by re-scanning 10% of samples.

The second mode of operation explored is *data reduction*. Temporal coding in structured light scanning often precludes motion while scanning is in progress. Spatial codes, which utilize color or pseudorandom dot distributions, enable all depth information to be captured in a single frame. However, these techniques do not lend themselves to tightly-packed high resolution samples [Salvi, et al. 2004]. Image frequency can be utilized to plan a sampling of the scene with optimally placed dot locations. Reconstruction quality is thus improved while maintaining sparse samples. While not implemented during this thesis, the use of dique coloring in high density regions with less-accurate, simple spatial pattern distortion in low density regions can greatly reduce the number of colors to be discriminated.

The key technology behind this structured light sensor is the DMD, a pixel-independent array of micromirrors that can be programmed to reflect or discard light. DMD arrays are approaching cameras in resolution and frame rate. DMDs are much higher resolution than beam sensors and even flash LIDAR arrays and more importantly, require no actuation. Future intelligent sensors are envisioned which couple low-resolution, low-throughput range receivers, micromirrors and co-located color cameras to determine optimal samples in the same vein as our sensor.

6.2.5 Discussion

An approach to image-directed, Lumenhanced scanning was described in this case study. The key idea is that image frequencies correlate with geometric frequency in constrained planetary environments. Thus, image information can be utilized as a prior in directing range scanning of a scene. Validation of the approach was conducted for lunar landing and ground mapping with simulation and terrestrial analog terrain respectively. Preliminary results are promising: reconstruction improvements of 40% can be achieved over naive scanning techniques from a single viewpoint. Simple, but effective visualization for sparse meshes was demonstrated by combining image-estimated texture with the geometric portion of a lunar surface BRDF. Lastly, a hardware implementation of a pixel-independent image-directed scanner was developed.

In the future, the effects of sensor noise on the approach should be quantitatively analyzed. Noise in both the scanner and the camera can contribute to inaccurate prediction and wasted samples. While

this problem was mitigated in these experiments by limiting the maximum regional density that could be sampled, the sensors were low noise to begin with. The ability to handle a larger number of different albedos could help extend this technique to related environments (such as the terrestrial underground). Possibilities to crosscheck albedo changes (which contribute to frequency content in images) and shape changes with multi-flash imagery and probabilistic albedo estimation techniques should be investigated. Lastly, while this approach is defeated with strong cast shadowing, these shadows can be detected and mitigated with a variety of active illumination.

6.3 Utilizing the Sun as a Motional Source

Planetary applications have the great advantage of the sun. While complex illumination resulting from interaction of the sun with atmosphere is a detriment to terrestrial application, this thesis has demonstrated that the simple conditions of direct sunlight on the moon enable superior image understanding. Rotation of cosmic bodies produces a span of appearances related by their geometry and distribution of illumination³⁰. From the reference of any surface point however, the sun is tantamount to a moving source. Recording changes in appearance over time can produce information and reduce uncertainty about the scene beyond any single instantaneous snapshot. This section discusses two planetary applications that can benefit from incorporating many solar-constrained images using surprisingly simple vision techniques and presents some preliminary results.

6.3.1 Super-Resolution Revisited

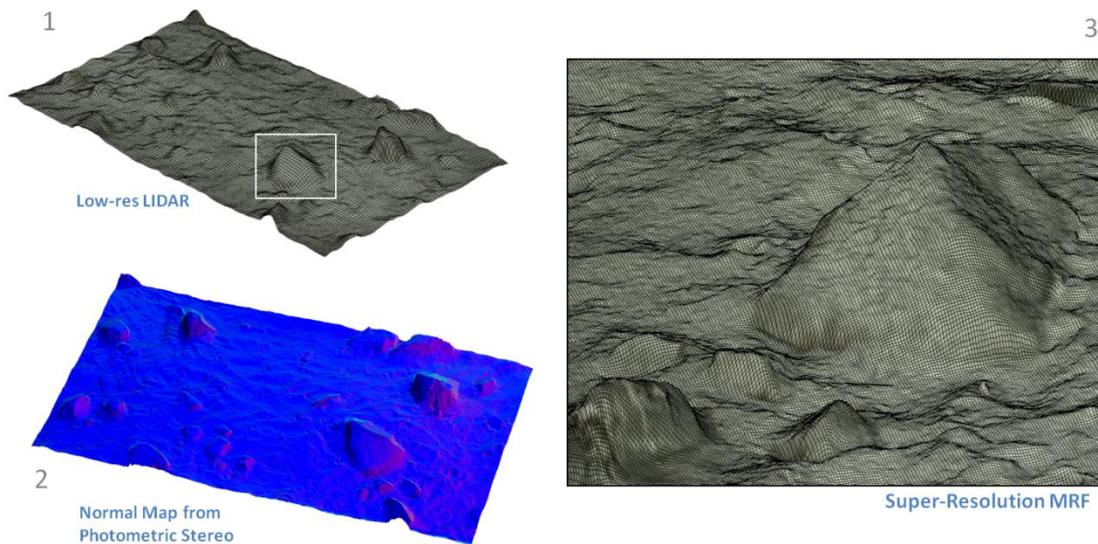


Figure 79. Many images of the scene taken from a single view point with moving sun can further enhance the super-resolution approach by utilizing photometric stereo. (1) A low resolution LIDAR model is fused with the shading estimates from multiple images using photometric stereo (2) to create a super-resolution model.

The MRF super-resolution method described in section 6.1 Lumenhances models by incorporating high resolution geometric information from a single constrained image. The key idea is that surface normals can be recovered directly from Lambertian planetary images, as pixel intensity scales with the derivative

³⁰ But not necessarily the angle or position of the source.

(gradient) of the geometry. This shape from shading (SFS) method is distinct from prior range fusion approaches which assume a linear relationship between image gradients and range gradients.

However, there are many problems with estimating shape from a single image. Variegation of materials is perhaps the most difficult to handle. Albedo must either be estimated in a separate phase (as proposed here) or optimized concurrently, which has the possibility for greater accuracy, but may fail catastrophically for difficult scenes such as the coal mine. Sensor noise can also have profound effects on numerical stability or the direction of shading. Noise was mitigated in the SFS experiments by taking HDR images³¹ and using gradient direction cues – specifically tilt angle – from the low resolution LIDAR model. Integrability is enforced by projecting the normals onto a Fourier basis [Frankot, et al 1988]. However, this method limits the improvement from images³² by destroying much of the high frequency information.

Surface normals can be recovered with much greater accuracy from sets of solar-constrained images using a simple process called *photometric stereo* (see Figure 80). If the scene reflectances are Lambertian, image intensity can be written as the dot product of the surface normal and the light vector, with a scalar albedo (6.33).

$$\begin{aligned} E &= \rho \hat{n} \cdot \vec{l} \\ E &= \vec{l} \cdot (\rho \hat{n}) \end{aligned} \quad (6.33)$$

The albedo can also be factored with the surface normal if unit vectors are not required. Exploiting this dot product relationship, equation (6.33) can be written in matrix form to describe image formation for all pixels in the scene under many lighting conditions (6.34):

$$\begin{array}{c} \updownarrow \text{images} \\ \begin{bmatrix} E_{1,1} & \dots & E_{1,k} \\ \vdots & & \vdots \\ E_{i,1} & \dots & E_{i,k} \end{bmatrix} = \begin{bmatrix} \vec{L}_1 \\ \vdots \\ \vec{L}_i \end{bmatrix} \times \left(\rho \begin{bmatrix} n_1 \\ \vdots \\ n_k \end{bmatrix} \right) \end{array} \quad (6.34)$$

← k pixels →

where $E_{i,k}$ is the intensity of the k th pixel in the i th image. The light source vectors corresponding to each image (L_i) are written as row vectors in the L matrix, while k normals are written as column

³¹ which reduces saturation and averages out salt and pepper noise

³² in an effort to limit *reductions* in accuracy which are much worse

vectors in the N matrix. Notice that this formulation only works under the condition of single, point illumination for each image. This linear system of equations can be solved to recover the normal matrix in the least squares sense by pseudo-inversion for 3 or more images (6.35).

$$L^{-1}E = \hat{N} \quad (6.35)$$

The albedo for each scene point can be recovered from the N matrix by normalizing all the vectors and taking the magnitude as the albedo. The implication of this formulation is that the rank of the the L matrix must be full, meaning that the sources must be non-coplanar on the unit sphere [Woodham 1980]. Moreover, the accuracy of reconstruction depends specifically on a wide angular separation of the source between images. While this does limit the approach in equatorial lunar terrain, for example, there is good possibility of applying photometric stereo to lunar polar regions or on tumbling asteroids.

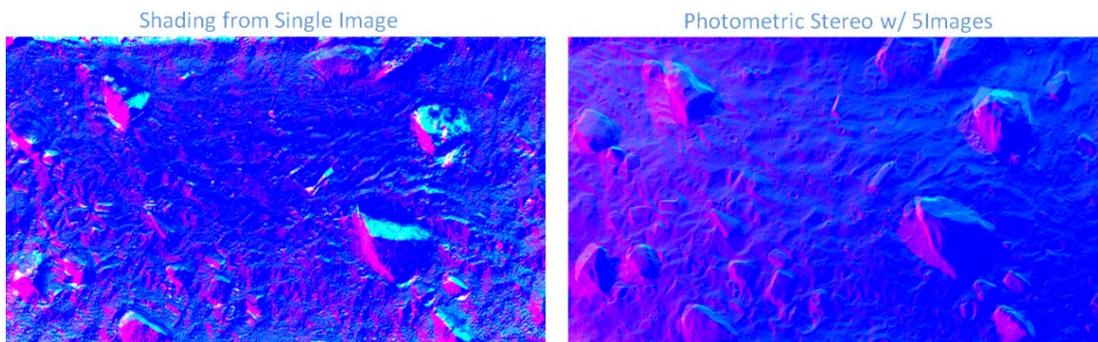


Figure 80. Normal maps of the moonyard estimated from intensity images. Estimating normals from many images using photometric stereo produces cleaner, more accurate normals than the single image technique. Moreover, seeding the tilt angles with sparse range information is no longer necessary.

Woodham’s photometric stereo was used to recover shape from the moonyard terrain for super-resolution Lumenhancement. This was used as a drop-in replacement for the SFS algorithm with a single image without changing any other part of the MRF or method. While newer nonlinear optimization approaches like [Goldman, et al. 2005] and [Alldrin, et al. 2008] can estimate the BRDF along with the illumination from several images (thus handling non-Lambertian scenes), these methods are not required for robust reconstruction in planetary terrains. Lambertian reflectance and solar direction are givens in the domain.

Figure 80 shows a comparison of the normal map generated using the stereo and SFS methods. The stereo method demonstrates less error and preservation of higher frequency detail, particular in the

quasi-flat regions. There is a source of error not present in the SFS method however. The equation assumes each pixel is lit directly from line of sight to the light source. If the images contain large areas of cast shadows, the accuracy of the data is affected particularly when using few images. Loss of detail can be seen in the moonyard data by observing the base of the large rocks, which are prone to shadowing.

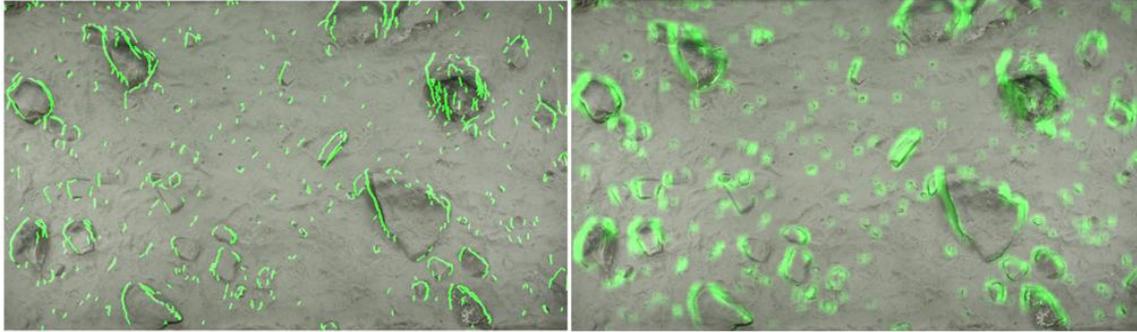


Figure 81. Shadow Maps can clearly identify shadows and occlusion edges (left). Detected edges (right) can be expanded to create a cost map in MRF fusion to discount for inaccurate interpolation data. Note that the color images shown are not the shadowed images used. These ambient images are used for visual clarity. The position of the sun goes from left to right.

Shadows can be explicitly detected by observing solar motion using the method of [Raskar, et al. 2004].

The key idea is that by comparing the pixel value in an image with the maximum value seen across all images under different lighting conditions a shadow ratio R_k can be created for each image of the form:

$$R_k = \frac{(\bar{l}_k \cdot n)}{\max_i (\bar{l}_i \cdot n)} \quad (6.36)$$

Shadows can then be detected in each image by simply finding pixels where the shadow ratio is close to 0³³. The caveat is that a region must be unshadowed in at least a single image for the shadow to be detected. However, if this not the case, there is no added detriment over utilizing raw data. Shadow maps of the moonyard were generated from the same set of solar imagery used to perform photometric stereo. Shadowed data was discard in photometric analysis with the method of [Forsyth, et al. 2002]. The occlusion edges³⁴ detected using Raskar's approach are used to create a cost map for the MRF that discounts for range interpolation in these areas and promotes the high resolution image data Figure 81. Using this approach it is shown that a 64x reconstruction (8x by 8x downsampling) was enhanced 42%

³³ It will likely not be zero due to interreflections.

³⁴ occlusions can be conveniently detected by traversing the source-camera epipolar lines looking for negative step edges.

over bilinear interpolation. The standard SFS method showed enhancement of 32%, meaning that the **multi-image method exhibited a 31% improvement over a single image.**

6.3.2 Material Classification using Gonioreflectometry In Situ

Lunar skylights are environments of immense interest for exploration, due to a variety of aforementioned reasons. Mobile mapping robots which inch up to the rim may be able to peer in laterally, identifying the existence of a lava tube. Extreme science robots may even rappel down to the floor to map and gather samples. Some of these applications are described in [Jones, et al. 2012], where some of the data is generously loaned for this work. Many dangers that planetary surface robots may encounter and anomalies that must be investigated stem from material properties, not just geometry. This section describes two skylight related applications which require identification of lunar materials from imagery. A method proposed to address these applications leverages possibility for high resolution orbital views of the skylight from satellite imagery over many illumination conditions.

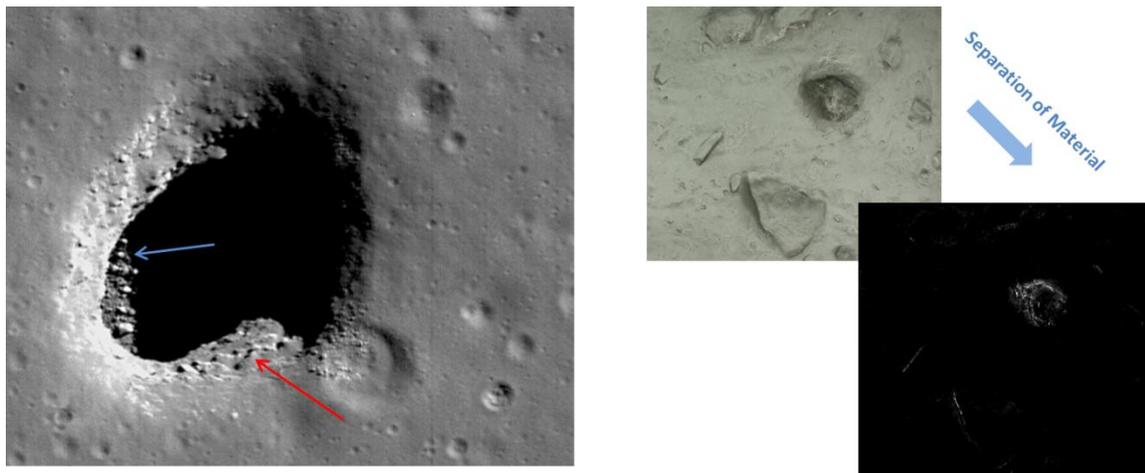


Figure 82. Materials can be identified with physics-based segmentation techniques. The Marius Hills skylight (left) has anomalous features that can inform the geology of the region. The image on the (right) shows a rock from the moonyard where regolith has brushed off due to the steep slopes. Lumen enhancement can identify that it is a different material than the background, while not being fooled by other regolith covered rocks – a task difficult even for a human.

Before sending a surface robot on a potentially destructive mission to explore a skylight rim, it must first be determined whether the rim is firm rock capable of supporting a robot's weight, loosely compacted regolith, or some combination of the two. Figure 82 shows an image of the Marius Hills skylight with a sloped, ragged rim (red arrow) consisting of bright ripples. It is unknown if exposed bedrock can be seen in these images, creating convenient "stepping stones" or whether the ripples are an artifact of deep regolith and illumination conditions, meaning that the edge may be dangerously slippery. While the angle of repose determines the maximum slope regolith can be fashioned by itself, this value varies with

the region and soil type; these materials cannot be discriminated from this single intensity image or even a geometric model of the skylight.

A second consideration exists for robots that will operate on the skylight floor. The blue arrow in Figure 82 shows what appears to be large boulders at the bottom of the skylight, lit by glancing sunlight. Likewise, it cannot be ascertained whether these artifacts are boulders, exposed bedrock or regolith that may present entrapment hazards. While boulders are a geometric feature, range scanning is unlikely to model the floor with certainty due to finite beam divergence and range error at orbital distances.

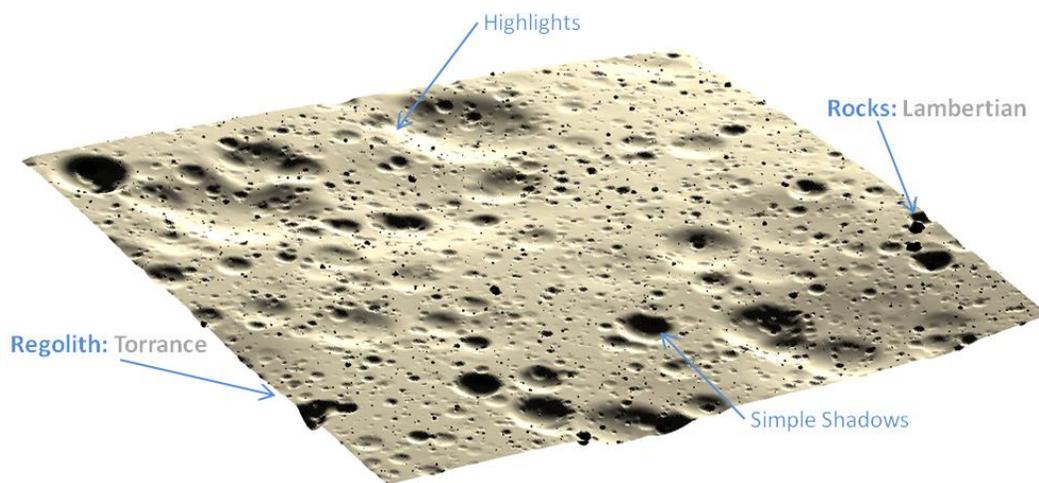


Figure 83. Simulated Rendering of Moon Terrain. Materials from this thesis are utilized to generate fidelity of appearance. The underlying regolith-covered terrain uses a Torrance-Sparrow model, while the rocks use a Lambertian model. The physical model is joint work documented in [Jones, et al. 2012].

A photometric stereo based approach can be used to discriminate materials assuming that they are actually discriminable with visible spectrum imagery. In fact, one of the stated objectives of the NASA LRO satellite is to infer shape from photometric methods³⁵. Only a small modification of the approach is necessary for added material information.

The matrix formulation of equation (6.34) implies albedos (i.e. materials) can be recovered directly from the magnitude of the surface normal vectors providing the scene is Lambertian. If the scene is non-lambertian, neither surface normal nor albedo can be recovered or separated with accuracy. However, this approach can still be used to test the “diffuseness” of a scene. Residuals from least-squares

³⁵ <http://lunar.gsfc.nasa.gov/lroc.html>, last visited April 2012.

estimation can be used to distinguish materials when only a few classes exist. Recall that the estimated normal matrix N is given by:

$$\begin{aligned} (\rho_{est} \hat{N}) &= (L^T L)^{-1} L^T E \\ (\rho_{est} \hat{N}) &= L^+ E \end{aligned} \quad (6.37)$$

where ρ_{est} are the estimated Lambertian albedos. These can be used to reproject the estimates back into the pixel intensity space to recover the error of estimation at each point (the residuals):

$$\hat{E} = L(\rho_{est} \hat{N}) \quad (6.38)$$

$$\varepsilon = E - \hat{E} \quad (6.39)$$

Analysis of the histogram of error distributions (ε) and the albedo map (ρ_{est}) can reveal natural classes of the data. Given the close reflectivity and the highly variegated nature of these materials, it is impossible to distinguish these materials from a single image. However differences in highlights and albedo are more significant across images with many illumination conditions. Thresholding of the histogram between the peaks of the distribution can be used for classification, or clustering can be used to create contiguous regions with robustness to noise.

This is particularly effective in the lunar case, where there is a single uncommon material (rocky surfaces) against a much more common background (regolith). If we further set the matrix ρ_{est} for all points to some scalar ρ_{mean} estimated from some known regolith pixels, the residuals become magnified by the slight differences in albedo (6.40), this is called the *material* metric.

$$\begin{aligned} score_{material} &= E - L(\rho_{mean} \hat{N}) \\ \rho_{mean} &= f(\rho_{est}) \end{aligned} \quad (6.40)$$

This simple algorithm is tested on simulated lunar terrain from an orbital view (Figure 83). The terrain model has two parts, a mesh model of the undulating ground with large craters and a sparse distribution of much smaller rocks using the standard Surveyor distribution. The terrain is rendered with the CMU-1 Torrance BRDF found from goniorelectometry and the rocks are rendered with the Lambertian granite model. The granite uses a false diffuse albedo very close to the regolith to test the robustness of the

approach (real granite is much lighter). Attached shadows are rendered but not cast shadows or interreflection. Images are orthographic and 1000x1000 pixel resolution for a 20m x 20m area.

A second set of data (of the same terrain) emulates a 5% salt and pepper noise from camera acquisition (Figure 84). As the rocks are so tiny and their appearance so similar to the numerous craters and terrain, a human would have great difficulty detecting even a few correctly in the example images. Depending on the angle and orientation of view, craters can also appear to be convex or concave as a result of shading ambiguity.

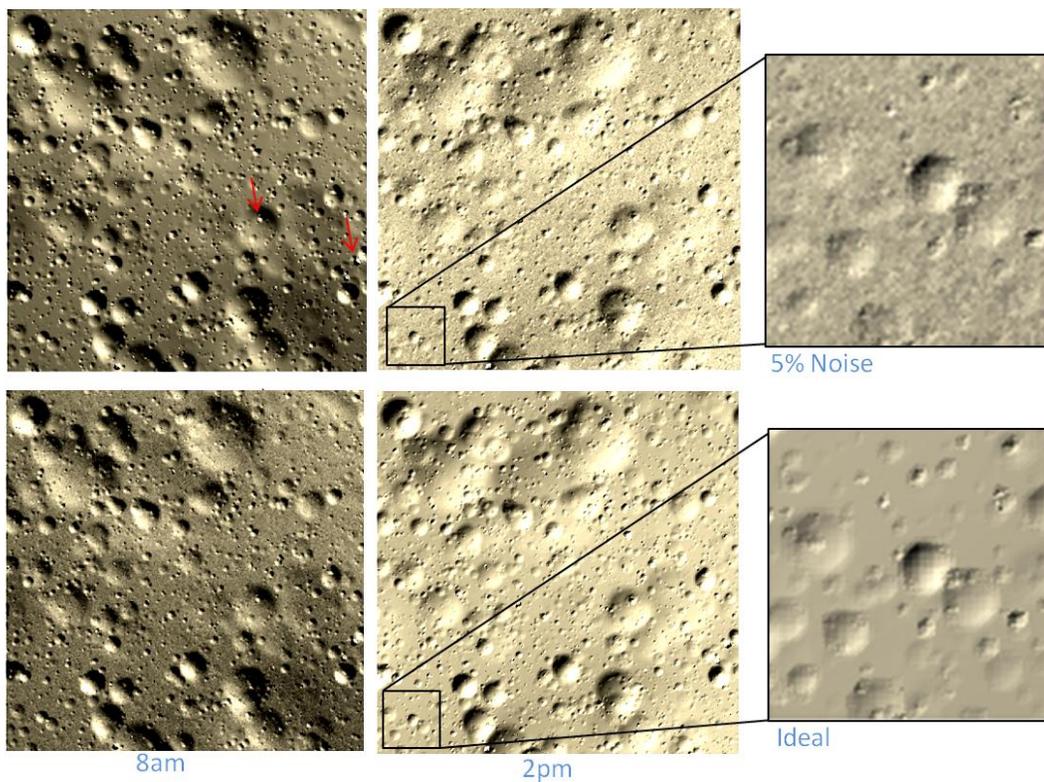


Figure 84. Simulated Orbital Imagery of the Moon from Two Sun Angles. The top row simulates the effect of pixel noise in the camera while the bottom images are ideal. Contrast is enhanced here for viewing. Even then, it is almost impossible to distinguish rocks in the images because they are small and have similar appearance to the regolith. Hint: red arrows point to some rocks which appear as bright blips.

This simulation is of course a simplification of lunar appearance; the very tops of rocks on the moon are likely covered with regolith making clusters of the material less consolidated and detection more difficult. However, there are still areas, particularly on the sides of rocks and crater rims where the rock is exposed. The simulated data suffices for the purposes of demonstrating the approach. Efficacy has

been further demonstrated on the physical moonyard data, where lunar variegation conditions are more accurate (see Figure 82 above).

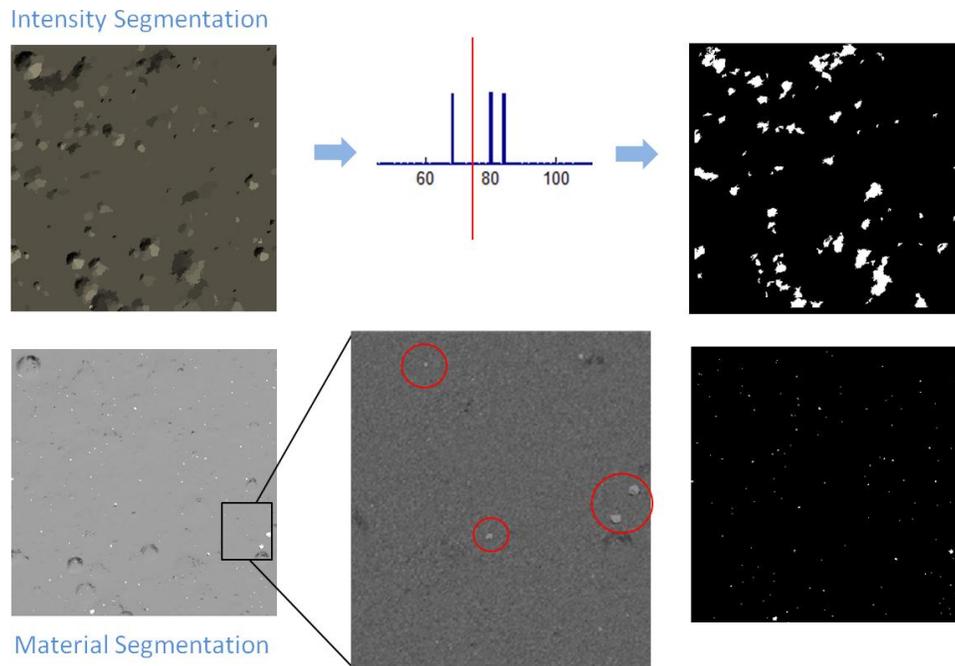


Figure 85. Comparison of Segmentation between a Naïve and the proposed Material Approach. Naïve segmentation uses an LAB transformation and thresholding, while the material approach uses photometric stereo from several images. Both algorithms use mean shift clustering.

Photometric material classification is compared against a naïve single-image approach. The single-image technique transforms the RGB data into the $L^*A^*B^*$ space and clusters similar pixels using mean shift. A final two-class separation is made by thresholding a histogram of the cluster centroids. In the multi-image approach, residuals from RGB color channels are merged with clustering and a histogram threshold is utilized to generate a classification in the same way. A total of 5 images are used. Figure 85 shows segmentation and classification using both approaches. Using only a single intensity image generates large clusters from shadowed regions and craters. Although rocks are a different material, they are simply merged into large clusters due to similarity of appearance. The material approach produces much better localization of feature, even in the presence of noise. In Figure 85, rocks are clearly shown as regions of high error.

Table 13. Performance of Algorithms in Material Classification

	Ideal + Material	5% Noise + Material	Ideal + Naïve Clustering
Total Rock Pixels	6283	“	“
Total Regolith Pixels	1194998	“	“
True Positive (tp)	1144	1144	2181
True Negative (tn)	1194998	1194980	982659
False Positive (fp)	0	18	212339
False Negative (fn)	5139	5139	4102
Accuracy	0.99	0.99	0.81
Precision	1	0.96	0.01

Table 13 summarizes the results of the experiment. There were a total of 254 rocks in the simulation, occupying 6283 pixels of the roughly 1 megapixel image. The approaches were used to identify which pixels were of the regolith and the rock material. The method of clustering before classifying can be seen as a prior over the contiguous nature of the material appearance. In the table above, true positives (tp) are pixels predicted as rock which are actually rock, while true negatives (tn) are pixels predicted as regolith which are actually regolith, and so on.

The results show that the multi-image method never misclassifies a rock as regolith (fp) in the ideal case, and there is only minimal misclassification in the case of noise. This attribute is advantageous as it can quickly narrow the search space of infrequent materials for manual human validation. Material segmentation is also relatively insensitive to noise, resulting in only 18 more pixels being classified incorrectly as rock. The single-image naïve approach was less accurate and much less precise. It predicted many regions were rock, resulting in an increase in both the number of true and false positives. As the total number of rock pixels was minuscule compared to the regolith, the large imprecise regions are not helpful to identification. Overall, the **multi-image method was about 22% more accurate and 100x more precise**. Equations for calculating the accuracy and precision of the classifiers are given below.

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} \quad (6.41)$$

$$precision = \frac{tp}{tp + fp} \quad (6.42)$$

It is perhaps more useful to look at spatial accuracy of the two classes. Clustering frequently underestimated the extent of rock regions, but predicted rock clusters were entirely contained within the area of real rocks, indicating high discriminativity. Figure 86 illustrates the number of rocks identified as a regional portion of actual rocks. Almost all of the clusters identified were within 50% of the correct size and shape of the actual rocks. Over half the rocks were identified with greater than 75% spatial accuracy. Noise affects clustering accuracy by about 10%.

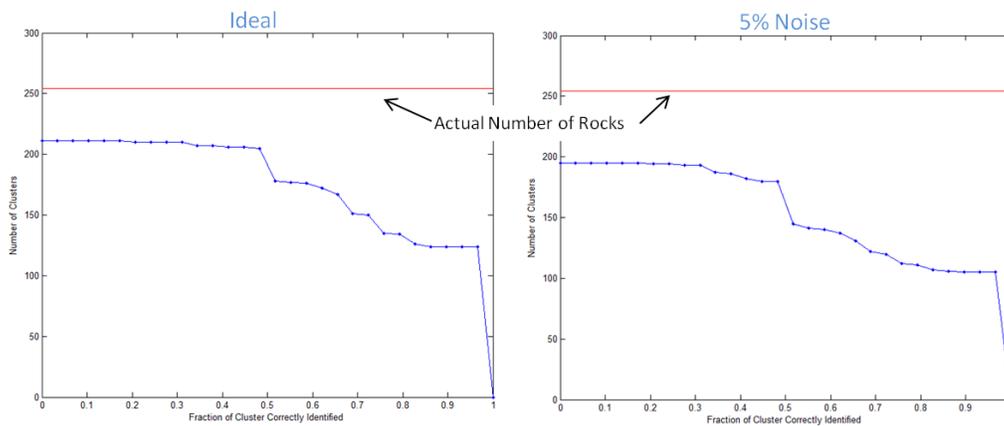


Figure 86. Number of Rocks Detected vs. Accuracy of the Cluster compared to the true region.

The approach can be further extended to handle many materials with greater accuracy. There is the possibility of stitching satellite imagery from multiple perspectives³⁶ using a sparse 3D model (the fusion concept) and parametric estimation of the BRDF at each point using nonlinear optimization. This approach is coined *gonioreflectometry in situ*. Though the exact implementation on lunar data is left to future work, the process is akin to that used for material gonioreflectometry documented in this thesis, except on a planetary scale. The problem can be further simplified with known sun direction and a sparse LIDAR model of the surface.

³⁶ Multi-perspective, multi-illuminant data exists for the Marius Hills hole. However, as of 2012 examples released to the public have lacked color and resolution capable for accurate material recovery.

Chapter 7:

Conclusion

This thesis introduces an innovative view of robot application domains as appearance spaces. These appearance spaces provide a convenient instrument for targeting and generalizing physics-based vision techniques to a spectrum of optically similar environments. By exploiting optical domain knowledge with active illumination and intensity imaging, the quality of geometric modeling is greatly improved – a process coined as Lumenhancement. The efficacy of Lumenhancement has been demonstrated experimentally in planetary spaces, a domain with highly advantageous, constrained appearance and relevant application.

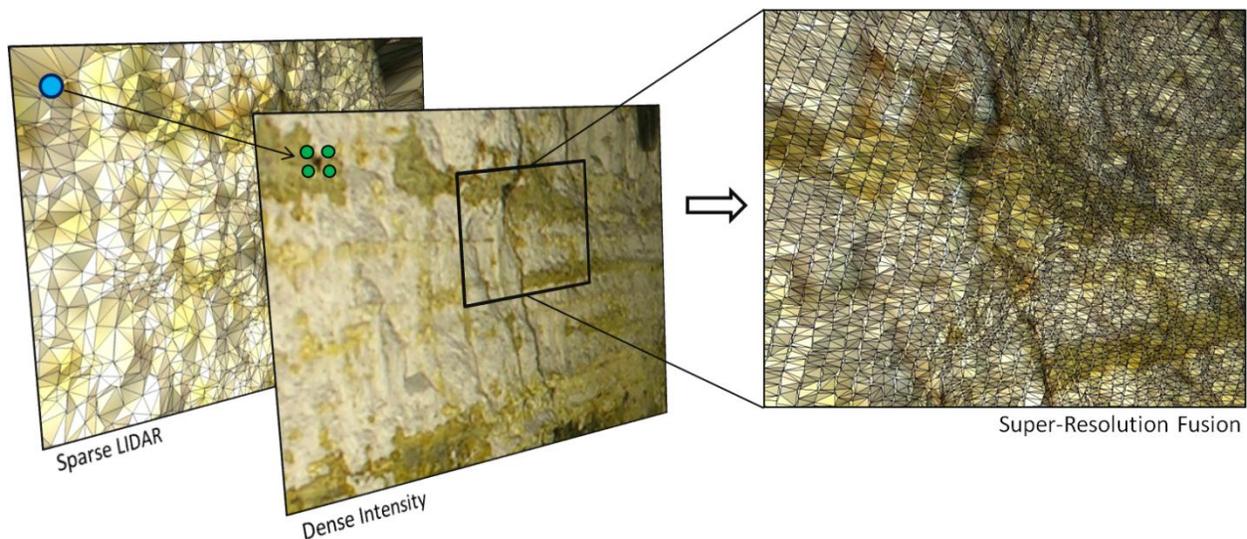


Figure 87. Super-resolution modeling of this mine wall demonstrates a generational leap of model quality using Lumenhancement.

The application of Lumenenhancement to planetary spaces has emphasized the need for reevaluation of common beliefs in robotic modeling. In particular, this thesis has shown that:

- Despite immense prior work in robotic modeling, significant improvements are still possible with targeted sensing.
- New perception techniques and sensor designs exploiting domain knowledge can bridge the gap between performance and generality.
- Previously unrelated environments can be associated by their similar appearance and application, enabling sharing of techniques.
- Proper environmental constraints can bring vision approaches out of the lab and into the field; classical algorithms can acquire renewed significance.

7.1 Summary of Results

This research characterized the geometric, material and atmospheric sensor physics of the planetary domain using exhaustive experimentation and statistical analysis. Utilizing this knowledge, novel methods were developed to enhance 3D models using active illumination, including geometric super-resolution, image-directed sampling and material classification. These methods were implemented on real robots and used data from actual sensors in analogous, physical planetary environments. Lastly, new visualization possibilities were explored to conveying these models to a human audience with increased effectiveness. A summary outline of important results follows:

Domains and Appearance Spaces

The idea of domains as appearance spaces spanning environmental geometric, material and illumination attributes was developed. Four constraining characteristics of planetary appearance were introduced: *barren*, *dry*, *rocky* and *simple illumination*. The simplifying effect of these assumptions on image formation was discussed along with the relationship between intensity images and geometry. Intra-domain variation in planetary appearance was analyzed by qualitative consideration of exemplary environments.

Planetary Characterization

Material and geometric aspects of the planetary domain were characterized through experimentation. A commodity gonireflectometer design was introduced, which enables rapid analysis of non-perfect and oversize field samples. The reflectance functions of seven planetary surface materials were recovered

using gonioreflectometry and contrasted with two ideal non-planetary materials. Planetary materials were found to be well-represented by simple Lambertian reflectance with low error. In situ analysis of material distributions furthermore showed that variegation is low and that even complex planetary spaces can be broken down into micro-environments of almost-uniform surface material. Lastly, analysis of representative surface and macroscopic geometry demonstrate the smoothness constraint is satisfied to a high degree.

Lumenhancement Case Studies

Case studies introduced three new, independent Lumenhancement techniques. First, fusion of LIDAR and camera data for super-resolution models was demonstrated for mine mapping with a mobile robot. Shape estimates were recovered from high resolution images take under artificial, point illumination and integrated with low density LIDAR readings in a Markov random field. Second, artifact-free models were acquired by directing LIDAR sampling of a scene using image frequency features. This technique was demonstrated for lunar landing application using simulated data as well as lunar surface mapping using an analog moonyard. A novel image-directed structured light sensor was developed to produce pixel-independent samplings of scenes utilizing the approach. Lastly, a method was developed to classify materials by recovering their BRDF through motion of the light source. This technique targets science autonomy applications in space exploration, such as identifying rocks.

New adaptations of non-photorealistic rendering methods were explored for model display. Point rendering for super-resolution clouds demonstrated enhanced geometric continuity for dense, but noisy model data. Adaptive surface frequency display showed that altering the roughness of material BRDFs according to image frequency can enhance the realism of data-deficient meshes without increase in storage or computation.

7.2 Contributions

The significant contributions of this thesis are:

- Innovation in the view of environments as constrained appearance spaces and the utilization of their properties to analyze and target sensing approaches.
- Characterization of the planetary domain for the geometric and material properties of appearance.

- Development of **Lumenhancement** which enables superior quality 3D modeling by exploiting these appearance properties.
- Creation of the **first empirical graphics BRDF database of planetary materials** utilizing gonioreflectometry. Further development and characterization of an optical lunar regolith simulant, CMU-1, enables accurate recreation of lunar appearance for future vision research.
- Creation of the **first expansive database of range sensor comparative performance** in any domain.
- Development of **three new algorithms for Lumenhancement** of resolution, sample selection and material classification. These techniques enhanced resolution by 40x, reduced reconstruction error by up to 40% and increased classification rates by 22% on experimental data.
- Development of **two novel sensors** including (1) an image-directed structured light scanner that combines several principles of Lumenhancement at the hardware level and (2) a commodity gonioreflectometer that provides reflectance data for difficult, oversized and non-ideal environmental samples.
- Demonstration of a class of non-photorealistic techniques – point rendering, adaptive surface frequency variation and displacement mapping – for visualizing Lumenhanced models with enhanced effectiveness.
- Creation of the **most accurate and extensive mine maps** to date (in Bruceton Mine) using Lumenhancement. Field experimentation also conducted first surveys of and generated first models in Walker’s Mill Cave. These models represent significant improvement in quality from prior work.

7.3 Impact

Contributions from this thesis extend the state-of-the-art in several ways. Firstly, future underground applications are anchored by quantification of the physical and material properties encountered. Analysis of the correctness of common modeling assumptions and planetary implications to intensity imaging advised the performance of techniques developed here and will advise future techniques.

Datasets generated in this work have provided crucial training data for humans and robots alike and performance benchmarks for progeny not previously available. The virtual reconstructions of Bruceton Mine are of historical as well as technical significance. Furthermore, study of environmental variation within the domain and correlation with neighboring domains will guide efficient reuse of successful applications across a variety of environments.

The innovations in data collection, interpretation and display presented herein denote generational leaps over prior work. These innovations represent a paradigm shift in targeting domain-specific sensing with environmental knowledge. Other outdoor robotic applications will similarly benefit from the type of domain considerations refined and advocated by this work. The plethora of planetary environments alone and the necessity of robotic operations in them promise broad technical significance for this work. The realism of field implementation stressed in the thesis further ensures specific implementation will remain practical and relevant for years to come.

The application focus of this work will generate renewed awareness of the problems for humans in planetary operations and emphasize the compelling motivation for robots. The great potential of these modeling and visualization techniques for improving underground reconnaissance, in particular, brings the ideal of practical, economical subterranean robotics one step closer to reality. Ultimately, this work will facilitate the proliferation of robots into these and other hazardous environments and similarly increase the safety of human workers.

Lastly, robotic modeling with Lumenenhancement will enable new remote scientific discoveries by providing unprecedented quality of data with immersive optical sensing. Lumenhanced geometric models will create the next MoonView™ or CaveView™ where future explorers can visit these remote worlds with the fidelity and virtualization expected of the human experience, but from the convenience of their computer.

7.4 Future Work

This thesis sets the groundwork for critical analysis and design for environmental appearance using Lumenenhancement. However further investigation is still possible and necessary in planetary and other domains. Recommendations for future work include:

Further characterization of planetary materials and spaces. This work conducted characterization of planetary appearance in several exemplary environments. While the characterization documented is the most extensive to date, it is by no means complete. Lava tubes for example, are environments of great interest for Lunar and terrestrial application, but they are completely ignored here. Additional characterization of geological properties in situ – such as surface weathering – and correlation to appearance are also of great interest to science applications. Continued research in this area will contribute to greater domain knowledge for exploitation.

Explicit evaluation of noise and algorithm sensitivity. Preliminary analysis of environmental variation in the planetary domain was conducted. Inferences were made regarding the effect of these variations on core assumptions of this work. Accuracy, error and modes of failure were discussed for each of the Lumenhancement techniques introduced. However, a thorough, theoretically grounded investigation of the effect of noise on these techniques was not conducted. Use of Lumenhancement would also benefit greatly from studies contrasting the relative performance of algorithms within the domain and the numerical significance of the planetary assumptions in bounding these errors.

Implementation of real-time sensors on mobile robots. A class of novel Lumenhanced sensors was proposed in this work and a single proof-of-concept design was presented. However, more work is required to for the realization these technologies on mobile robots. Algorithms and hardware sensors should be married in common, real-time packages conducive to strap-down integration. The relative performance of these sensors should also be compared against traditional approaches, much like the range accuracy study detailed here, to inform conditions of use and future sensor designs.

Extension of Lumenhancement to other advantageous environments and domains. Lumenhancement was demonstrated for planetary environments; however related environments like the seafloor and atmospheric surfaces (i.e. Mars) were also discussed. Techniques proposed in this thesis could apply to these spaces with additional considerations for atmospheric scattering and exploitation of physics based sensing such as polarization and multiple sources. Ultimately, the purpose of Lumenhancement is to generalize optical sensing in appearance-constrained spaces. Thus, conclusive demonstration of its efficacy is successful utilization in many other domains, such as indoor, aerial and urban robotics.

References

1. **[Achanta, et al. 2009]** Achanta, R., Hemami, S., Estrada, F., Susstrunk, S. Frequency-tuned Saliient Region Detection. In Proc. CVPR, 2009.
2. **[Achtuthan, et al. 2010]** A. Achuthan, M. Rajeswari, D. Ramachandram, M. Aziz, I. Shuaib. Wavelet energy-guided level set-based active contour: A segmentation method to segment highly similar regions. *Computers in Biology and Medicine*: 40, 2010. Pp 608-620.
3. **[Agarwal, et al. 2006]** A. Agarwal, R. Raskar, R. Chellappa. What is the Range of Surface Reconstructions from a Gradient Field? In Proc. ECCV 2006.
4. **[Albee, et al. 2001]** A. Albee, R. Arvidson, F. Palluconi, T. Thorpe. Overview of the Mars Global Surveyor Mission. *Journal of Geophysical Research*: 106: 2001.
5. **[Aldroubi, et al. 2001]** A. Aldroubi, K. Grochenig. Nonuniform Sampling and Reconstruction in Shift-Invariant Spaces. *SIAM Review*: 43(4), 2001.
6. **[Alldrin, et al. 2008]** N. Alldrin, T. Zickler, D. Kriegman. Photometric Stereo with Non-Parametric and Spatially-Varying Reflectance. In CVPR, 2008.
7. **[Anguelov, et al. 2010]** D. Anguelov, C. Dulong, D. Filip, C. Frueh, et al. Google Street View: Capturing the World at Street Level. In *Computer*: 43(6), 2010. Pp. 32-38.
8. **[Archinal, et al. 2006]** B. Archinal, M. Rosiek, R. Kirk, B. Redding. The Unified Lunar Control Network 2005. US Geological Survey, Open-File Report 2006-1367.
9. **[Ashikhmin, et al. 2000]** M. Ashikhmin, P. Shirley. An Anisotropic Phong BRDF Model. *Journal of Graphics Tools*: 5(2), 2000. Pp. 25-32.
10. **[Baltzakis 2003]** H. Baltzakis, A. Argyros, P. Trahanias. Fusion of Laser and Visual Data for Robot Motion Planning and Collision Avoidance. *Machine Vision and Applications*, 2003.
11. **[Banta 1996]** J. Banta. M. Abidi. Autonomous placement of a range sensor for acquisition of optimal 3-D models. Proc. IEEE IECON, 1996.
12. **[Baumgart 1975]** B. Baumgart. A Polyhedron Representation for Computer Vision. National Computer Conf. AFIPS, 1975.
13. **[Bell, et al. 2003]** J. Bell, S. Squyres. K. Herkenhogg, J. Maki, H. Arneson, D. Brown, et al. Mars Exploration Rover Athena Panoramic Camera (Pancam) investigation. *Journal of Geophysical Research*: 108(E12), 2003.
14. **[Blake 1985]** A. Blake. Boundary conditions for lightness computation in the Mondrian World. In *Computer Vision, Graphics and Image Processing*: 32(3), 1985. Pp. 314-327.
15. **[Bouguet 2001]** J. Bouguet. Camera Calibration Toolbox for Matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/index.html. 2001.
16. **[Boulanger, et al. 2001]** P. Boulanger, J.-F. Lapointe. Creation of a Live Virtual Reality Model of Mining Environments from Sensor Fusion. In Proc. Minespace 2001.
17. **[Braquelair, et al. 2005]** A. Braquelair, B. Kerautret. Reconstruction of Lambertian Surfaces by Discrete Equal Height Contours and Regions Propagation. *Image and Vision Computing*, 2005. Volume 23(2), pp. 177-189.
18. **[Cassinelli, et al. 2010]** Cassinelli, A., Zerroug, A., Ishikawa, M. Camera-less Smart Laser Projector. ACM SIGGRAPH 2010 Emerging Technologies.
19. **[Chambers, et al. 2011]** A. Chambers, S. Achar, S. Nuske, J. Rehder, B. Kitt, et al. Perception for a River Mapping Robot. In Proc. IROS 2011.
20. **[Champeny-Bares, et al. 1991]** L. Champeny-Bares, L. Coppersmith, K. Dowling. The terregator mobile robot. Carnegie Mellon University, Robotics Institute Technical Report, 1991. CMU-RI-TR-93-03.
21. **[Chang, et al. 2001]** I. Chang, R. H. Park. Segmentation based on Fusion of Range and Intensity Images using Robust Trimmed Methods. *Pattern Recognition* 2001.

22. **[Cheng, et al. 2005]** Y. Cheng, A. Johnson, L. Matthies. MER-DIMES: A planetary landing application of computer vision. In Proc. CVPR, 2006.
23. **[Cheng, et al. 2006]** Y. Cheng, M. Maimone, L. Matthies. Visual Odometry on the Mars Exploration Rovers. IEEE Robotics and Automation Magazine Special Issue (MER): 13(2), 2006.
24. **[Christoudias, et al. 2002]** C. Christoudias, B. Georgescu, P. Meer. Synergism in Low Level Vision. In Proc. ICPR, 2002.
25. **[Clements 2003]** M. Clements. Shotcreting in Australian Underground Mines: A Decade of Rapid Improvement. In Shotcrete Spring 2003.
26. **[Cord, et al. 2003]** A. Cord, P. Pinet, Y. Daydou, S. Chevrel. Planetary regolith surface analogs: optimized determination of Hapke parameters using multi-angular spectro-imaging laboratory data. Icarus: 165(2), 2003. pp 414-427.
27. **[Criswell 1972]** D. Criswell. Lunar Dust Motion. Proc. Third Lunar Science Conference, 1972.
28. **[Dana, et al. 1999]** K. Dana, B. Van-Ginneken, S. Nayar, J. Koenderink. Reflectance and Texture of Real World Surfaces. In ACM Transactions on Graphics: 18(1), 1999.
29. **[Daubechies 1992]** Daubechies, I. Ten Lectures on Wavelets. Soc. for Industrial and Applied Mathematics, 1992.
30. **[De Angelis, et al. 2002]** G. De Angelis, J. Wilson, M. Cloudsley, J. Nealy, et al. Lunar Lava Tube Radiation Safety Analysis. Journal of Radiation Research: 43, 2002.
31. **[Debevec 1998]** P. Debevec. Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography. In Proc. ACM SIGGRAPH, 1998.
32. **[Debevec, et al. 1997]** P. Debevec, J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. In Proc. SIGGRAPH, 1997.
33. **[Diebel, et al. 2005]** J. Diebel, S. Thrun. An Application of Markov Random Fields to Range Sensing. Neural Information Processing Systems, 2005.
34. **[Dunlop, et al. 2007]** Dunlop, H. Thompson, D., Wettergreen, D. Multi-scale Features for Detection and Segmentation of Rocks in Mars Images. In Proc. CVPR, 2007.
35. **[Edwards, et al. 2006]** L. Edwards, M. Broxton. Automated 3D Surface Reconstruction from Orbital Imagery. Proc. AIAA Space, 2006.
36. **[Fairfield, et al. 2005]** N. Fairfield, G. Kantor, D. Wettergreen. Three Dimensional Evidence Grids for SLAM in Complex Underwater Environments. Proc. International Symposium of Unmanned Untethered Submersible Technology, 2005.
37. **[Ferrie, et al. 1989]** F. Ferrie, M. Levine. Where and Why Local Shading Analysis Works. Proc. IEEE PAMI, 1989.
38. **[Finlayson, et al. 2001]** G. Finlayson, S. Hordley, P. Hubel. Color by Correlation: A Simple, Unifying Framework for Color Constancy. Proc. IEEE PAMI, 2001.
39. **[Fong, et al. 2008]** T. Fong, M. Bualat, M. Deans, M. Allan, et al. Field testing of utility robots for lunar surface operations. In Proc. AIAA Space, 2008.
40. **[Forsyth, et al. 2002]** D. Forsyth, J. Ponce. Computer Vision: A Modern Approach. Prentice Hall, 1st Ed. 2002.
41. **[Fraden 2003]** J. Fraden. Handbook of Modern Sensors: Physics, Designs, and Applications. Springer, 3rd edition, 2003.
42. **[Francisco, et al. 2011]** F. Francisco, O. Bertolami, P. Gil, J. Paramos. Modelling the reflective thermal contribution to the acceleration of the Pioneer spacecraft. Preprint, arXiv:1103.5222, 2011.
43. **[Frankot, et al. 1988]** R. Frankot, R. Chellappa. A Method for Enforcing Integrability in Shape from Shading Algorithms. IEEE PAMI: 10(4), 1988.
44. **[Fruh, et al. 2003]** C. Fruh, A. Zakhor. Data Processing Algorithms for Generating Textured 3D Building Façade Meshes from Laser Scans and Camera Images. Proc. 3D Data Processing Visualization and Transmission, 2003.
45. **[Fujii, et al. 2009]** Fujii, K, Grossberg, M., Shree, N. A Projector-Camera System with Real-Time Photometric Adaptation for Dynamic Environments. In Proc. CVPR, 2009.

46. **[Goldberg, et al. 2002]** S. Goldberg, M. Maimone, L. Matthies. Stereo Vision and Rover Navigation Software for Planetary Exploration. In Proc. IEEE Aerospace Conference, 2002.
47. **[Goldman, et al. 2005]** D. Goldman, B. Curless, A. Hertzmann, S. Seitz. Shape and Spatially-Varying BRDFs from Photometric Stereo. In Proc. ICCV 2005.
48. **[Gooch, et al. 1998]** A. Gooch., B. Gooch., P. Shirley, E. Cohen. A Non-photorealistic Lighting Model for Automatic Technical Illustration. Proc. ACM SIGGRAPH, 1998.
49. **[Gooch, et al. 2001]** B. Gooch, A. Gooch. Non-photorealistic Rendering. A K Peters Ltd., 2001.
50. **[Gould, et al. 2008]** S. Gould, P. Baumstarck, M. Quigley, et al. Integrating Visual and Range Data for Robotic Object Detection. In Proc. ECCV 2008.
51. **[Grossman, et al. 1998]** J. Grossman, W. Dally. Point Sample Rendering. In Rendering Techniques 1998.
52. **[Gu, et al. 2002]** X. Gu, S. Gortler, H. Hoppe. Geometry Images. ACM Trans. on Graphics: 21(3), 2002.
53. **[Gupta, et al. 2008]** M. Gupta, S. Narasimhan, Y. Schechner. On Controlling Light Transport in Poor Visibility Environments. Proc IEEE CVPR, 2008.
54. **[Hapke 1993]** B. Hapke, R. Nelson, W. Smythe. The Opposition Effect of the Moon – The Contribution of Coherent Backscatter. In *Science*, 1993.
55. **[Hapke 1998]** B. Hapke, R. Nelson, W. Smythe. The Opposition Effect of the Moon: Coherent Backscatter and Shadow Hiding. In *Icarus*, Volume 133, Issue 1, 1998.
56. **[Harrison, et al. 2009]** A. Harrison, P. Newman. Image and Sparse Laser Fusion for Dense Scene Reconstruction. Proc. Field and Service Robotics, 2009.
57. **[Haruyama, et al. 2009]** J. Haruyama, et al. Possible lunar lava tube skylight observed by SELENE cameras. Geophysical Research Letters, 2009.
58. [Hawkins, et al. 2007] S. Hawkins, J. Boldt, E. Darlington, R. Espiritu, R. Gold, B. Gotwols, M. Grey, C. Hash, J. Hayes, S. Jaskulek, et al. The Mercury Dual Imaging System on the MESSENGER spacecraft. Space Science Reviews: 131, 2007. Pp 247-338.
59. **[Haya, et al. 2004]** Hata, S., Sumioka, K. Detection of obstacles on the moon. IEEE Symposium on Industrial Electronics, 2004.
60. **[Hayakawa 1994]** K. Hayakawa. Photometric Stereo under a Light Source with Arbitrary Motion. Journal of the Optical Society of America, 1994.
61. **[Heiken, et al. 1991]** Heiken, G., Vanniman, D., French, B. The Lunar Sourcebook. 1991.
62. [Helfenstein, et al. 1987] P. Helfenstein, J. Veverka. Photometric Properties of Lunar Terrains Derived from Hapke's Equation. *Icarus*: 72, 1987. pp 342-357.
63. **[Holte, et al. 2008]** M. Holte, T. Moeslund, P. Fihl. Fusion of Range and Intensity Information for View Invariant Gesture Recognition. Proc. IEEE CVPR, 2008.
64. **[Horau, et al. 1988]** On the Geometric Interpretation of Image Contours. In Artificial Intelligence, 1988.
65. **[Horn 1970]** B. K. P. Horn. Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View. MIT Artificial Intelligence Lab, Technical Report 232, 1970.
66. **[Horn 1986]** B. K. P. Horn. Robot Vision. The MIT Press, 1986.
67. **[Hussman, et al. 2008]** S. Hussman, T. Ringbeck, B. Hagebeucker. A Performance Review of 3D TOF Vision Systems in Comparison to Stereo Vision Systems. In book Stereo Vision, edited by A. Bhatti. 2008.
68. **[Ikeuchi, et al. 1979]** K. Ikeuchi, B. K. P. Horn. An Application of the Photometric Stereo Method. MIT Artificial Intelligence Lab, AI Memo No. 539, 1979.
69. **[Izadi, et al. 2011]** S. Izadi, D. Kim, O. Hiliges, D. Molyneaux, R. Newcombe, P. Kohli, et al. Kinectfusion: Real-time 3D Reconstruction and Interaction using a Moving Depth Camera. ACM Symposium on User Interface Software and Technology, 2011.
70. **[Jensen, et al. 2001]** H. Jensen, S. Marschner, M. Levoy, P. Hanrahan. A Practical Model for Subsurface Light Transport. In SIGGRAPH 2001.
71. **[Johnson, et al. 1996]** A. Johnson, M. Hebert. Seafloor Map Generation for Autonomous Underwater Vehicle Navigation.
72. **[Johnson, et al. 2008]** J. Johnson, M. Shepard, W. Grundy. Spectrogoniometric Measurements and Models of Lunar Analog Soils. Lunar and Planetary Science, 2008.

73. **[Jones, et al. 2012]** Jones, H. Wong U., Peterson, K., Koenig, et al. Complementary Flyover and Rover Sensing for Superior Modeling of Planetary Features. Submitted for publication, 2012.
74. **[Kajiya 1986]** J. Kajiya. The rendering equation. In Proc. SIGGRAPH, 1986.
75. **[Kennedy 2009]** H. Kennedy. Introduction to 3D Data: Modeling with ArcGIS 3D Analyst and Google Earth. Wiley & Sons, 2009.
76. **[Kirk, et al. 2007]** R. Kirk, E. Howington-Kraus, M. Rosiek, et al. Ultrahigh Resolution Topographic Mapping of Mars with HiRISE Stereo Images: Methods and First Results. In Proc. ISPRS, 2007.
77. **[Landau 1967]** H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. Acta Math: 117, 1967. Pp. 37-52.
78. **[Lee, et al. 1985]** C. H. Lee, A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. In *Artificial Intelligence*, 1985.
79. **[Lee, et al. 1990]** H. Lee, E. Breneman, C. Schulte. Modeling Light Reflection for Computer Color Vision. Proc. IEEE PAMI, 1990.
80. **[Lee, et al. 1998]** G. Chungmou Lee, G. Stockman. Detection of Object Wings in Fused Range and Intensity Imagery. Pattern Recognition 1998.
81. **[Levoy, et al. 2000]** M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz et al. The Digital Michelangelo Project: 3D scanning of Large Statues. Proc. ACM SIGGRAPH 2000.
82. **[Li 2001]** S. Li. Markov Random Field Modeling in Image Analysis. 2001.
83. **[Liang, et al. 1996]** S. Liang, J. Townshend. A Modified Hapke Model for Soil Bidirectional Reflectance. Remote Sensing of Environments: 55(1), 1996.
84. **[Lide 1994]** D. Lide. "Characteristics of Particles and Particle Dispersoids." Handbook of Chemistry and Physics, 75th Edition. CRC Press 1994.
85. **[Lindner, et al. 2008]** M. Lindner, M. Lambers, A. Kolb. Sub-pixel data Fusion and Edge-enhanced Distance Refinement for 2D/3D images. International Journal of Intelligent Systems Technologies and Applications, 2008.
86. **[Loupias, et al. 2000]** Loupias, E., Sebe, N., Bres, S., Jolion, J. Wavelet-based Salient Points for Image Retrieval. In Proc. 4th International Conference on Advances in Visual Information Systems, 2000.
87. **[Maimone, et al. 2006]** M. Maimone, J. Biesiadecki, E. Tunstel, Y. Cheng, et al. Surface Navigation and mobility intelligence on the Mars Exploration Rovers. Intelligence for Space Robotics, TSI Press, 2006. Ch. 3.
88. **[Malick 1987]** J. Malik. Interpreting Line Drawing of Curved Objects. International Journal of Computer Vision, 1987.
89. **[Mallick, et al. 2005]** S. Mallick, T. Zickler, D. Kriegman, P. Belhumeur. Beyond Lambert: Reconstructing Specular Surfaces Using Color. CVPR 2005.
90. **[Mark, et al. 2010]** Mark, L. H., Okouneva, G., Saint-Cyr, P., Ignakov, D., English, C. Near Optimal Selection of Views and Surface Regions for ICP Pose Estimation. In Adv. in Visual Computing, 2010.
91. **[Marschner 1999]** S. Marschner, S. Westin, E. Lafortune, et al. Image-Based BRDF Measurement Including Human Skin. Eurographics Workshop on Rendering, 1999.
92. **[Martinez, et al. 2007]** J. Martinez, A. Reina, A. Mandow. Spherical Laser Point Sampling with Application to 3D Scene Genetic Registration. Proc. IEEE ICRA, 2007.
93. **[Martinez, et al. 2007]** Martinez, J., Reina, A., Mandow, A. Spherical Laser Point Sampling with Application to 3D Scene Genetic Registration. In Proc. ICRA, 2007.
94. **[Mastin, et al. 2009]** A. Mastin, J. Kepner, J. Fisher. Automatic Registration of LIDAR and Optical Images of Urban Scenes. Proc. IEEE CVPR, 2009.
95. **[Matusik, et al. 2003]** W. Matusik, H. Pfister, M. Brand, L. McMillan. A Data-Driven Reflectance Model. Proc. ACM SIGGRAPH, 2003.
96. **[Matusik, et al. 2003]** W. Matusik, H. Pfister, M. Brand, L. McMillan. Efficient Isotropic BRDF Measurement. Eurographics Workshop on Rendering, 2003.
97. **[Maxwell, et al. 2008]** B. Maxwell, R. Friedhoff, C. Smith. A Bi-illuminant Dichromatic Reflection Model for Understanding Images. Proc. IEEE CVPR, 2008.

98. **[McKay, et al. 1994]** D. McKay, J. Carter, W. Boles, C. Allen, J. Allton. JSC-1: A New Lunar Soil Simulant. Engineering, Construction, and Operations in Space IV. American Society of Civil Engineers, 1994. pp 857-866.
99. **[Morris 2007]** A. Morris. Robotic Introspection for Exploration and Mapping of Subterranean Environments. PhD Dissertation. Robotics Institute, Carnegie Mellon University, 2007.
100. **[Morris, et al. 2006]** A. Morris, D. Ferguson, et al. Recent Developments in Subterranean Robotics. Journal of Field Robotics, Vol. 23, No. 1, Jan 2006, pp. 35-57.
101. **[Morris, et al. 2007]** A. Morris, U. Wong, Z. Omohundro, et al. 3D Modeling of Subterranean Environments by Robotic Survey. Carnegie Mellon University, Robotics Institute Tech Report, 2007.
102. **[Mostafa, et al 1999]** M. Mostafa, S. Yamany, A. Farag. Integrating Shape from Shading and Range Data. Proc. IEEE CVPR, 1999.
103. **[Mustard, et al. 1989]** J. Mustard, C. Pieters. Photometric Phase Functions of Common Geologic Minerals and Applications to Quantitative Analysis of Mineral Mixture Reflectance Spectra. Journal of Geophysical Research: 90(B10), 1989. pp 13619-13634.
104. **[Nagatani, et al. 2009]** K. Nagatani, T. Matsuzawa, K. Yoshida. Scan-point Planning and 3-D Map Building for a 3-D Laser Range Scanner in an Outdoor Environment. Proc. Field and Service Robotics, 2009.
105. **[Narasimhan 2002]** S. Narasimhan, S. Nayar. Vision and the Atmosphere. International Journal of Computer Vision, 2002.
106. **[Narasimhan, et al. 2003]** S. Narasimhan, V. Ramesh, S. Nayar. A Class of Photometric Invariants: Separating Material from Shape and Illumination. Proc. IEEE ICCV, 2003.
107. **[Nayar, et al. 1990]** S. Nayar, K. Ikeuchi, T. Kanade. Shape from Interreflections. Carnegie Mellon University, Robotics Institute Technical Report, 1990. CMU-RI-TR-90-14.
108. **[Nefian, et al. 2009]** A. Nefian, K. Hussmann, M. Broxton, et al. A Bayesian Formulation for Sub-pixel Refinement in Stereo Orbital Imagery. Proc. IEEE ICIP, 2009.
109. **[Ng, et al. 2005]** R. Ng, M. Levoy, M. Bredif, et al. Light Field Photography with a Hand-Held Plenoptic Camera. Stanford Tech Report CTSR 2005-02, 2005.
110. **[Oliver, et al. 2011]** A. Oliver, X. Xiong, B. Akinci, D. Huber. Automatic Creation of Semantically Rich 3D Building Models from Laser Scanner Data. In Proc. International Symposium on Automation and Robotics in Construction, 2011.
111. **[Omohundro 2007]** Z. Omohundro. Robot Configuration for Subterranean Modeling. PhD Dissertation. Robotics Institute, Carnegie Mellon University, 2007.
112. **[Oren, et al. 1994]** Oren, M., Nayar, S. K. Generalization of Lambert's Reflectance Model. SIGGRAPH, 1994.
113. **[Pedersen, et al. 2008]** L. Pedersen, C. Han, M. Vitus. Dark Navigation: Sensing and Rover Navigation in Permanently Shadowed Lunar Craters. International Symposium on AI, Robotics and Automation in Space, 2008.
114. **[Pfister, et al. 2000]** H. Pfister, M. Zwicker, J. Baar, M. Gross. Surfels: Surface Elements as Rendering Primitives. In Proc. SIGGRAPH, 2000.
115. **[Pugacheva, et al. 2005]** S. Pugacheva, V. Shevchenko. The parameters involved in Hapke's model for estimation of the composition of the ejecta lunar terrains. Sternberg State Astronomical Institute, Moscow University, 2005.
116. **[Pulli, et al. 1997]** K. Pulli, M. Cohen, T. Duchamp, et al. View-based Rendering: Visualizing Real Objects from Scanned Range and Color Data. Eurographics Rendering Workshop, 1997.
117. **[Rankin, et al. 2005]** A. Rankin, A. Huertas, L. Matthies. Evaluation of Stereo Vision Obstacle Detection Algorithms for Off-Road Autonomous Navigation. AUVSI Symposium on Unmanned Systems, 2005.
118. **[Raskar, et al. 2004]** R. Raskar, K. Tan, R. Feris, J. Yu, M. Turk. Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering using Multi-Flash Imaging. Proc. ACM SIGGRAPH 2004.
119. **[Ray, et al. 2010]** C. Ray, S. Reis, S. Sen, J. O'Dell. JSC-1A lunar soil simulant: Characterization, glass formation, and selected glass properties. Journal of Non-Crystalline Solids: 365, 2010. pp 2369-2374.
120. **[Reed 1997]** M. Reed, P. Allen, I. Stamos. Automated Model Acquisition from Range Images with View Planning. Proc. IEEE CVPR, 1997.

121. **[Rizzoni 2002]** Rizzoni, G. Principles and Applications of Electrical Engineering. 4th Ed., 2002. Pp 744.
122. **[Rosso, et al. 2006]** O. Rosso, M. Martin, A. Figiola, K. Keller, A. Plastino. EEG analysis using wavelet-based information tools. *Journal of Neuroscience Methods*: 153(2), 2006. Pp. 163-182.
123. **[Russell 1916]** H. N. Russell. On the Albedo of the Planets and Their Satellites. *Proc. National Academy of Science USA*: 2(2), 1916. pp 74-77.
124. **[Salvi, et al. 2004]** Salvi, J., Pages, J., Battle, J. Pattern Codification Strategies in Structured Light Systems. In *Pattern Recognition*, 2004. Pp 827-849.
125. **[Sato, et al. 2012]** H. Sato, B. Denevi, M. Robinson, B. Hapke, A. McEwen. Photometric Parameter Maps of the Moon from LROC WAC Observations. 43 Lunar and Planetary Science Conference, 2012.
126. **[Shafer 1985]** S. A. Shafer. Using color to separate reflection components. In *Color Research Applications*, 10:210-218, 1985.
127. **[Shahid, et al. 2007]** Shahid, K., Galina, O. Intelligent LIDAR scanning region selection for satellite pose estimation. *Computer Vision and Image Understanding*, Vol. 107, Issue 3, Sept. 2007. Pp 203-209.
128. **[Shenk, et al. 2002]** T. Schenk, B. Csatho. Fusion of LIDAR Data and Aerial Imagery for a More Complete Surface Description. *International Archives of Photogrammetry, Remote Sensing and Spatial Information*, 2002.
129. **[Shepard, et al. 2007]** M. Shepard, P. Helfenstein. A test of the Hapke Photometric Model. *Journal of Geophysical Research*: 112, 2007.
130. **[Smith, et al. 2006]** W. Smith, E. Hancock. Estimating Cast Shadows using SFS and Class-based Surface Completion. *Proc. ICPR*, 2006.
131. **[Smith, et al. 2010]** D. Smith, M. Zuber, G. Jackson, J. Cavanaugh, G. Neumann, H. Riris, et al. The Lunar Orbiter Laser Altimeter Investigation on the Lunar Reconnaissance Orbiter Mission. *Space Science Reviews*: 150, 2010. Pp 209-241.
132. **[Stein, et al. 2006]** A. Stein, A. Heurtas, L. Matthies. Attenuating Stereo Pixel-locking via Affine Window Adaptation. *Proc. IEEE ICRA*, 2006.
133. **[Stipes, et al. 2008]** J. Stipes, J. G. P. Cole, J. Humphreys. 4D Scan Registration with the SR-3000 LIDAR. *Proc. IEEE ICRA*, 2008.
134. **[Stork 2006]** D. Stork. Mathematical foundations for quantifying shape, shading and cast shadows in realist master drawings and paintings. *Proc. SPIE*, 2006.
135. **[Summers, et al. 2005]** V. Summers, A. Normoyle, R. Flo. Increasing Situational Awareness by Combining Realistic and Non-Realistic Rendering Techniques. *International Command and Control Research and Technology Symposium*, 2005.
136. **[Tardos, et al. 2002]** J. Tardos, J. Neira, P. Newman, J. Leonard. Robust mapping and localization in indoor environments using sonar data. *International Journal of Robotics Research*: 22, 2002. Pp. 311-330.
137. **[Thompson, et al. 2008]** Thompson, D., Smith, T., Wettergreen, D. Information-Optimal Selective Data Return for Autonomous Science and Survey. In *Proc. ICRA 2008*.
138. **[Thrun, et al. 2000]** S. Thrun, W. Burgard, D. Fox. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping. In *Proc. ICRA*, 2000.
139. **[Thrun, et al. 2005]** S. Thrun, S. Thayer, W. Whittaker, et al. Software Architecture of an Autonomous Robotic System. *IEEE Robotics and Automation Magazine*, 2004.
140. **[Thrun, et al. 2006]** S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, et al. Stanley: The Robot that Won the DARPA Grand Challenge. *Journal of Field Robotics* 2006.
141. **[Todt, et al. 2005]** S. Todt, C. Salama, A. Kolb. Real Time Fusion of Range and Light Field Images. *Proc. ACM SIGGRAPH*, 2005.
142. **[Torres-Mendez, et al. 2008]** L. Torres-Mendez, G. Dudek. Inter-Image Statistics for 3D Environment Modeling. *IJCV* (2008) 79. pp 137-158.
143. **[Urmson, et al. 2007]** C. Urmson, J. Anhalt, J. Bagnell, C. Baker, et al. Tartan Racing: A Multi-Modal Approach to the DARPA Urban Challenge. *Carnegie Mellon University Technical Report*, 2007.
144. **[Vandapel, et al. 1999]** N. Vandapel, S. Moorehead, W. Whittaker, R. Chatila, R. Murrieta-Gid. Preliminary Results on the Use of Stereo, Color Cameras, and Laser Sensor in Antarctica. *International Symposium on Experimental Robotics*, 1999.

145. **[Veeraraghavan, et al. 2007]** A. Veeraraghavan, R. Raskar, A. Agrawal, et al. Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing. Proc. ACM SIGGRAPH, 2007.
146. **[Venkataramani, et al. 2001]** R. Venkataramani, Y. Bresler. Optimal Sub-Nyquist Nonuniform Sampling and Reconstruction for Multiband Signals. IEEE Transactions on Signal Processing: 49(10), 2001.
147. **[Vineet, et al. 2009]** V. Vineet, P. J. Narayanan. Solving Multilabel MRFs using Incremental Alpha Expansion on the GPU. IIIT-Hyderabad, Tech Report, 2009.
148. **[Vogel, et al. 2009]** O. Vogel, M. Breuss, T. Leichtweis, J. Weickert. Fast Shape from Shading for Phong-Type Surfaces. In Scale Space and Variational Methods in Computer Vision. Lecture Notes in Computer Science: 5567, 2009.
149. **[Waggershauer 2005]** A. Waggershauer. Combining Full Spherical Depth and HDR Images to Implement a Virtual Camera. ISPRS Panoramic Photogrammetry Workshop 2005.
150. **[Walther, et al. 2006]** Walther, D., Koch, C. Modeling attention to salient proto-objects. Neural Networks, 2006.
151. **[Wang, et al. 2007]** Wang, Z., Boesch, R., Ginzler, C., Color and lidar data fusion: application to automatic forest boundary delineation in aerial images. Int. Arch. of the Photogrammetry Remote Sensing and Spatial Information Sciences, 2007.
152. **[Ward 1992]** G. Ward. Measurement and Modeling Anisotropic Reflection. Proc. ACM SIGGRAPH, 1992.
153. **[Wei, et al. 2006]** Wei, H., Bartels, M. Unsupervised Segmentation Using Gabor Wavelets and Statistical Features in LIDAR Data Analysis. In Proc. Pattern Recognition (ICPR), 2006.
154. **[Wen, et al. 2004]** J. Wen, H. Lu, W. Zhao. A Study of Truncated Cone-beam Sampling Strategies for 3D Mammography. IEEE Nuclear Science Symposium Conference Record, 2003.
155. **[Wettergreen, et al. 2005]** D. Wettergreen, P. Tompkins, C. Urmson, et al. Sun-Synchronous Robotic Exploration: Technical Description and Field Experimentation. International Journal of Robotics Research, Vol. 24, No. 1, 2005.
156. **[Wettergreen, et al. 2009]** D. Wettergreen, D. Jonak, D. Kohanbash, S. Moreland, S. Spiker, J. Teza. Field Experiments in Mobility and Navigation with a Lunar Rover Prototype. In Proc. Field and Service Robots, 2009.
157. **[Whyte, et al. 1996]** H. Whyte, D. Rye, E. Nebot. Localisation of automatic guided vehicles. In Robotics Research (ISRR), 1995.
158. **[Wilburn, et al. 2002]** B. Wilburn, M. Smulski, H. Lee, M. Horowitz. The Light Field Video Camera. Proc. Media Processors 2002, SPIE Electronic Imaging 2002.
159. **[Winnemoller 2006]** H. Winnemoller. Perceptually-motivated Non-Photorealistic Graphics. PhD Dissertation. Northwestern University, 2006.
160. **[Wohler, et al. 2008]** C. Wohler, P. d'Angelo. Stereo Image Analysis of Non-Lambertian Surfaces. International Journal of Computer Vision, 2008.
161. **[Wong, et al. 2009]** U. Wong, B. Garney, W. Whittaker, R. Whittaker. Camera and LIDAR Fusion for Mapping of Actively Illuminated Subterranean Voids. Proc. Field and Service Robotics, 2009.
162. **[Wong, et al. 2011]** U. Wong, A. Morris, C. Lea, J. Lee, C. Whittaker, B. Garney, W. Whittaker. Comparative Evaluation of Range Sensing Technologies for Underground Void Modeling. In Proc. IROS, 2011.
163. **[Woodham 1980]** R. Woodham. Photometric method for determining surface orientation from multiple images. Proc. SPIE, 1979. Vol. 155.
164. **[Worthington 2005]** P. Worthington. Re-illuminating single images using Albedo estimation. In Pattern Recognition, 2005.
165. **[Yamauchi 1997]** B. Yamauchi. A Frontier-based Approach for Autonomous Exploration. Proc. IEEE CIRA, 1997.
166. **[Yang, et al. 1997]** J. Yang, D. Zhang, N. Ohnishi, N. Sugie. Determining a Polyhedral Shape using Interreflections. Proc. IEEE CVPR, 1997.
167. **[Yang, et al. 2007]** Q. Yang, R. Yang, J. Davis, D. Nister. Spatial-Depth Super Resolution for Range Images. In Proc. IEEE CVPR, 2007.

168. **[Yoerger, et al. 1999]** D. Yoerger, A. Bradley, M. Cormier, et al. High Resolution Mapping of A Fast Spreading Mid Ocean Ridge with the Autonomous Benthic Explorer. Proc. International Symposium on Unmanned Untethered Submersible Technology, 1999
169. **[Zhang, et al. 1999]** R. Zhang, P. S. Tsai, J. E. Cryer, M. Shah. Shape from Shading: A Survey. Proc. IEEE PAMI, 1999.
170. **[Zhang, et al. 1999]** R. Zhang, P. Tsai, J. Cryer, M. Shah. Shape from Shading: A Survey. IEEE PAMI Vol. 21, No. 8, August 1999.
171. **[Zickler 2008]** T. Zickler. Color Subspaces as Photometric Invariants. International Journal of Computer Vision, 2008.
172. **[Zickler, et al. 2006]** T. Zickler, S. Mallick, D. Kriegman, P. Belhumeur. Color Subspaces as Photometric Invariants. Proc. IEEE CVPR, 2006.
173. **[Zuber, et al. 2000]** M. Zuber, D. Smith, A. Cheng, J. Garvin, et al. The Shape of 433 Eros from the NEAR-Shoemaker Laser Rangefinder. In Science: 289(5487), 2000. Pp 2097-2101.

Appendix

7.4.1 BRDF Polar Plots

Polar plot slices of fitted BRDFs are included in this section for completeness. These plots illustrate a constant-azimuth plot of the magnitude of radiant emergence (L_o , relative to unit Spectralon normal incidence given by distance from origin) as a function of view elevation angle (polar angle ϕ_o). Incident angles (ϕ_i) are modulated at 30, 60 and 90 degrees, indicated by the red vectors, to give 3 slices per BRDF. The full BRDF is of course continuous on the incident angle (ϕ_i), and three dimensional.

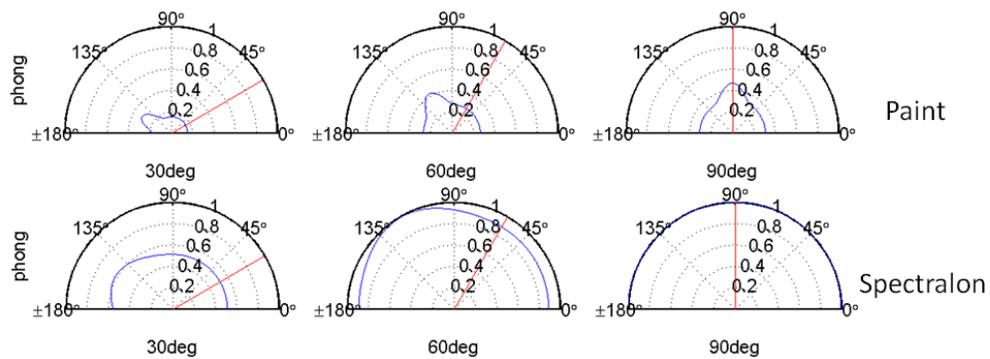


Figure 88. Reference Material BRDF Polar Plots.

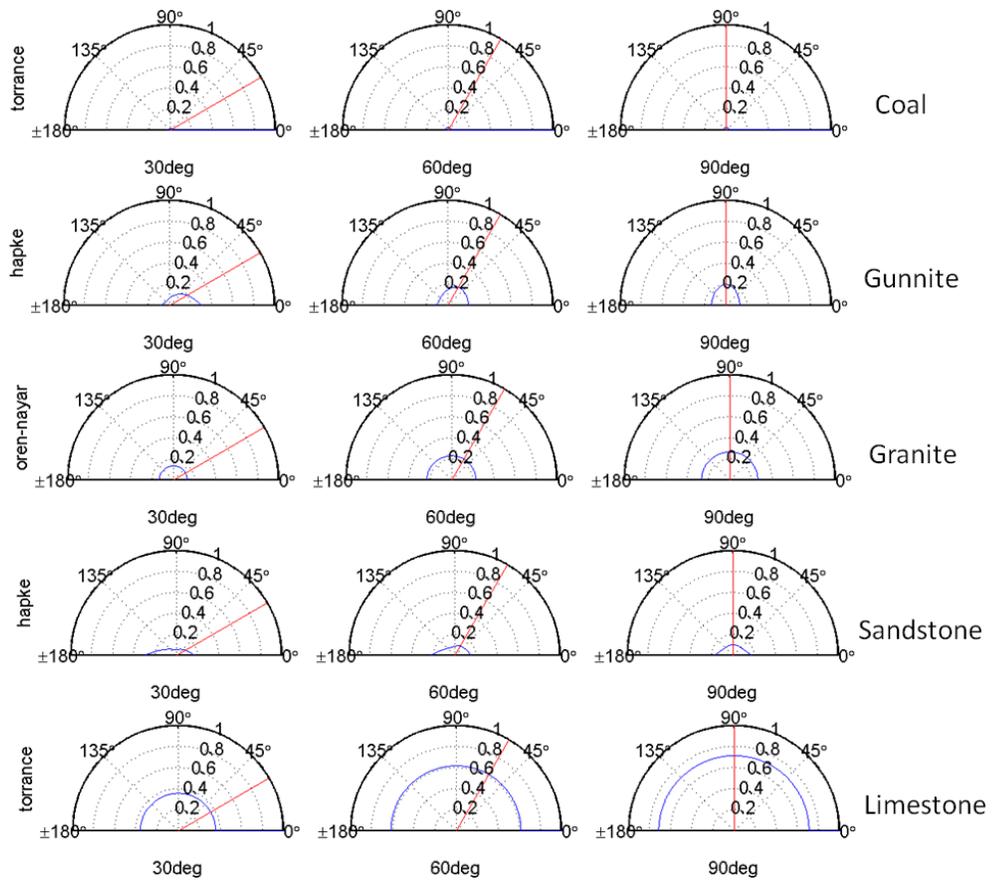


Figure 89. Underground Material BRDF Polar Plots.

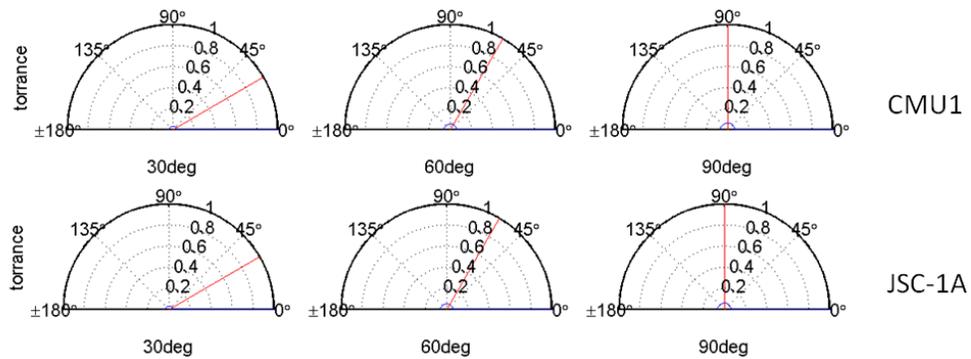


Figure 90. Planetary Material BRDF Polar Plots.