POLAR Traverse Dataset

(Polar Optical Lunar Analog Reconstruction)

User Documentation V1.0

November 2023

Developed at the Intelligent Robotics Group at NASA Ames

Research Center through the work of:

Margaret Hansen (CMU)

Dr. Uland Wong (NASA)

Dr. Terrence Fong (NASA)

This work was supported by NASA Space Technology Graduate Research Opportunity (NSTGRO) grant 80NSSC22K1206. The Solar System Exploration Research Virtual Institute (SSERVI) has graciously provided access to their laboratory facilities for collecting this data.

We would additionally like to thank Dr. Massimo Vespignani, Roshan Kalghatgi, and Molly O'Connor for their assistance with hardware setup, camera calibration, and terrain formation, respectively.

Contents

1 Description

The POLAR Traverse Dataset provides stereo pairs of camera images designed in such a way to mimic the traverse of a rover or other mobile platform across the lunar surface in polar regions. The polar regions of the moon are of interest to upcoming missions such as the Volatiles Investigating Polar Exploration Rover (VIPER) and the Artemis missions for lunar volatiles prospecting [1, 2] and are characterized by characterized by long shadows paired with areas of bright, direct sunlight due to their low solar incidence angles. Along with the reflectance properties of lunar regolith and the lack of atmospheric scattering to distribute light into shadows, this results in scenes with high variation in light intensity. This dataset was inspired by the POLAR Stereo Dataset, which is similar and provides images of a wide variety of such scenes on a smaller scale with no simulated camera motion. We believe the dataset will be useful for developing software algorithms for perception, mapping, and visual odometry under the harsh lighting conditions found in lunar south polar terrain.

In order to mimic a traverse across lunar terrain, a stereo pair of cameras was mounted on the end of a camera jib and suspended over the surface of a test bed containing a synthetic lunar scene. The camera jib was then rolled along the length of the test bed, with image pairs captured in 1 m increments. These images are taken at multiple exposure levels under two lighting conditions based on the light's incidence angle and moving in two directions along the test bed (facing "forward" or "reverse"). Additional variables that differ between traverses are the height and pitch of the cameras relative to the terrain. Dense ground truth is provided from a high-resolution LiDAR scanner, and estimated poses are provided from COLMAP [3]. This README details the setup, dataset organization, and file formats, with an appendix covering details on how COLMAP was used to generate poses.

The full dataset contains a total of 3,960 stereo pairs and is 13.4 GB (compressed). Individual "view" datasets are provided that contain the data for one lighting angle and one camera direction along the test bed - these are 3.4 GB max (compressed) and contain 990 stereo pairs each. The individual datasets contain poses and camera calibrations but do not contain the dense LiDAR scans; a dataset consisting only of these dense ground truth scans is available separately for download, and is 533.4 MB (compressed).

For questions, please contact Margaret Hansen at margareh@andrew.cmu.edu or Terry Fong at terry.fong@nasa.gov.

2 Experimental Setup

The POLAR Traverse Dataset consists of 6 traverses each across 4 views. The views are differentiated based on the viewing direction of the cameras along the test bed in addition to the azimuth angle of the light used to light the scene.[1] The viewing directions correspond to the x axis of the ground truth as discussed in Section 4.4 and shown in Figure 4 in that section (i.e. "forward" is facing in the direction of the positive x axis and "reverse" is facing in the direction of the negative x axis).

Two lighting azimuth angles were used during testing, 20° and 150°, both measured relative to the viewing direction to ensure comparability between the views with different viewing directions. Table 1 lists the 4 views for which data is recorded, along with the numeric labels used to refer to these views in the dataset,

[1] The terrain geometry does not change between the views in the dataset.

while Figure 1 shows the image collection and lighting positions for each viewing direction as arrayed along the test bed.

View	Viewing Direction	Lighting Angle
1	Forward	20°
2	Forward	150°
3	Reverse	20°
4	Reverse	150°

Table 1: Parameters for each view.

Traverse	Camera Height	Camera Pitch
1	1.3 m	14°
2	1.3 m	25°
3	1.3 m	35°
4	0.65 m	14°
5	0.65 m	25°
6	0.65 m	35°

Table 2: Parameters for each traverse.

The 6 traverses differ based on parameters related to the camera locations, namely the camera height above the terrain and the camera pitch. Three camera pitch values are available based on different static camera mounts used to secure the stereo pair to the jib arm: 14°, 25°, and 35°. Two height values, 1.3 m and 0.65 m, were used, with height measured relative to the ground outside of the test bed. The camera height and pitch associated with each traverse, along with the numeric labels used to refer to the traverses in the dataset, are listed in Table 2. Data is recorded for all 6 traverses for each of the 4 views.

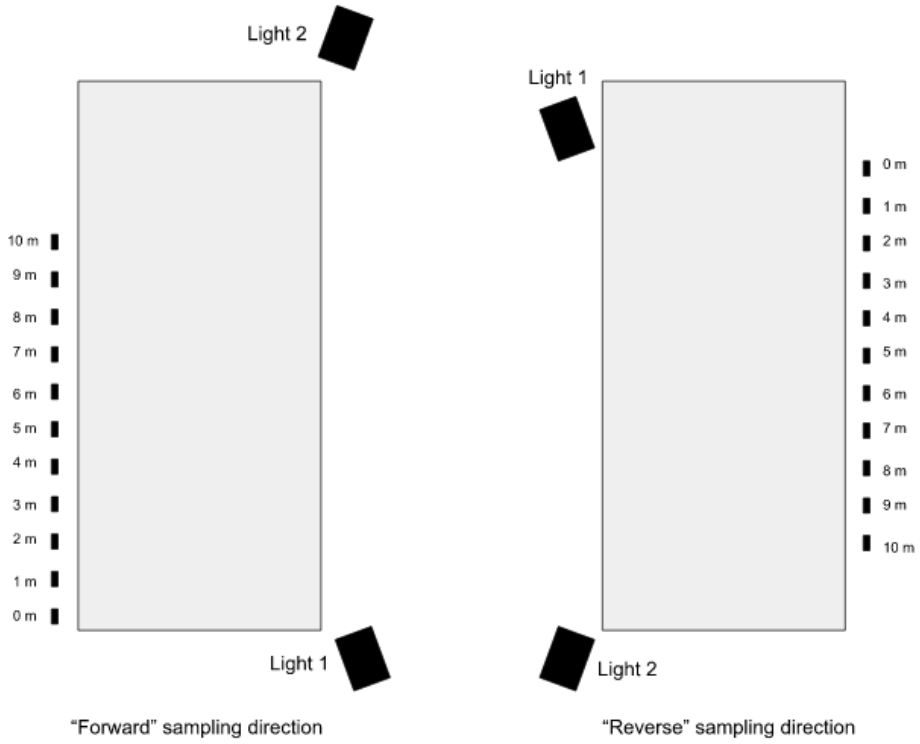


Figure 1: Diagrams showing approximate image collection and light positions for the two viewing directions. Light 1 corresponds to 20°and light 2 corresponds to 150°.

2.1 Scene Construction

The test bed is 18.6 m by 4 m and has been provided by the Solar System Exploration Research Virtual Institute (SSERVI). The terrain scene was constructed from lunar regolith simulant on the interior of the

test bed to reflect planetary science knowledge about the polar regions of the moon. There are a few craters of varying diameter, as well as uneven terrain and a few larger mounds in the scene. In the first two views, a human-made structure can be seen in the background of the images (also visible in Figure 2), included to simulate the presence of human-made structures on the surface. A light was placed at the corners of the test bed based on the desired lighting angles. The rest of the scene was blacked out to the extent possible to remove any external sources of light or reflection.

2.1.1 Regolith Simulant

The regolith simulant used in the test bed is a modified version of LHS-1 by Exolith Labs. LHS-1 is a general purpose simulant with relevant albedo for the lunar highlands that exhibits many of the same bulk optical phenomena as lunar regolith, such as the opposition effect, since it approximates a region of the particle size distribution of true regolith [4]. The modified version, mLHS-1, uses the same anorthosite and basaltic glass components of LHS-1 but leaves out the trace minerals for cost reasons due to the need to produce it in bulk quantities. mLHS-1 differs from LHS-1 less than 1% chemically and much less than 1% optically.

2.1.2 Scene Fidelity

This scene is designed to approximate stereo pair images from traversing lunar highlands terrain in polar regions to a high fidelity. However, there are areas in which it falls short. First, no effort was made to crop the images to the terrain in the test bed, resulting in images in which walls and other background materials are visible. In addition, due to the relatively close placement of the light source to the test bed, light fall-off according to the inverse square law can be observed across the scene, meaning the ideal exposure will be different for each position along the test bed. For actual lunar terrain, the light source (the sun) is far enough away that incoming light rays can be treated as parallel and motion on the scale of 10s of meters will not result in a noticeable difference between the angle of incidence of the light. Finally, the cameras were stationary during image collection, meaning there is no motion blur as would be expected for images taken while a platform is in motion.

2.2 Data Collection

2.2.1 Images

Two Allied Vision Manta G-419 cameras [5] paired with two Tamron M111FM08 8mm f/1.8 lenses [6], set at f8, were used to collect the images. The cameras were mounted to a stereo bar at a baseline of 40 cm. This stereo bar was then attached to a pitch mount to provide the desired pitch angle and finally attached to the end of a Proaim 10' Wave-2 jib crane [7]. This allows for the cameras to be suspended over the terrain and rolled along the length of the test bed for collecting images. The full setup, minus a pitch mount, is shown in Figure 2. In this example, the cameras are near the 1 m position for the "forward" facing direction along the test bed.

The pitch mounts were 3D-printed mounts designed to connect the stereo bar to the end of the camera jib arm such that a specific camera pitch is generated. Figure 3 provides an up-close view of the stereo bar, cameras, and a pitch mount attached to the end of the jib arm. Mounts were printed with three different pitch angles to simulate being able to change the pitch of the cameras.

Prior to data collection, the stereo camera pair was calibrated using a 9 x 12 checkerboard with 60 mm

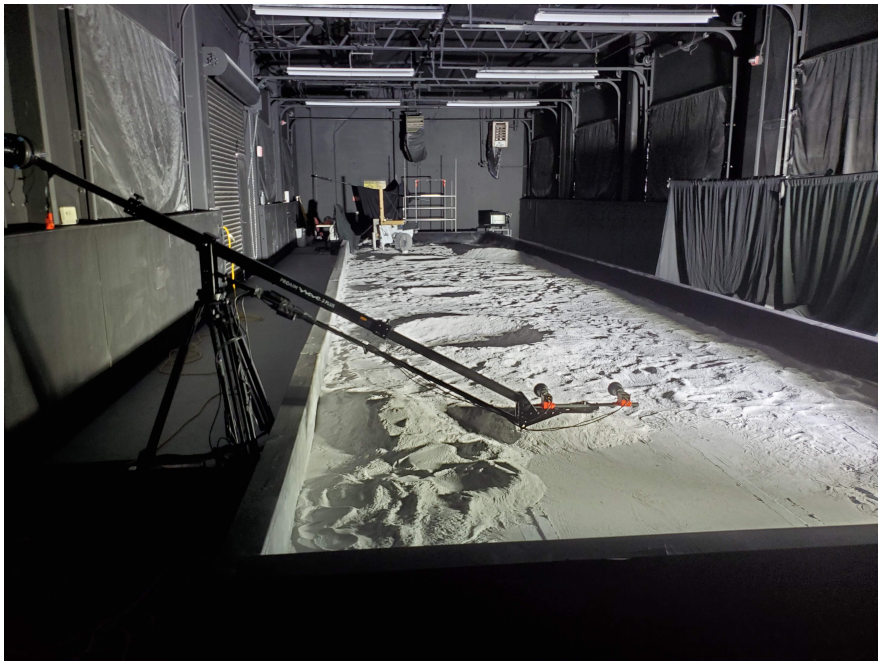


Figure 2: Camera jib arm with stereo bar and cameras suspended over test bed. Note that the setup shown in this image does not include a pitch mount.

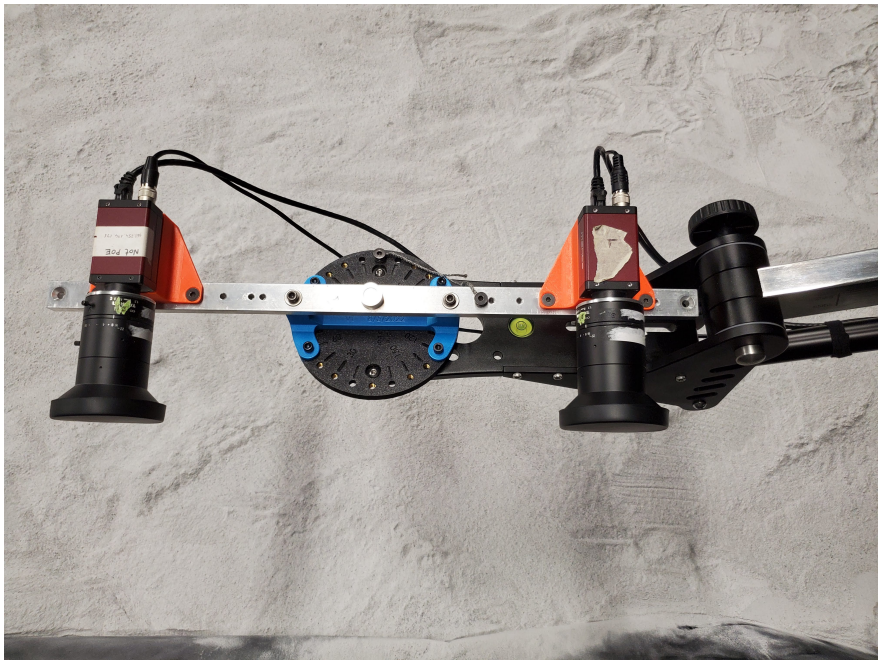


Figure 3: Close-up view of stereo bar mounted on jib arm with pitch mount (blue).

squares. These images were collected with overhead lighting and provided to MATLAB's Stereo Camera Calibrator App [8]. 8 image pairs were removed during the calibration due to missed detections of checkerboard corners, and 1 image pair was removed as an outlier in terms of reprojection error. The resulting calibration has an average reprojection error of 0.111 and is provided as part of the dataset.

During data collection, sets of stereo pair images differing by exposure time were captured for locations along the length of the test bed. The following 15 exposure times were used for each location (all in ms): 1, 5, 10, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500. 11 locations along the test bed were used, spaced out by approximately 1 m, with the dataset containing images from 0 m to 10 m, inclusive.

2.2.2 Ground Truth

Four ground truth scans of the terrain were collected using a FARO scanner in order to provide an accurate view of the scene. The approximate positions at which the scans were taken, along with the resulting composite scan, are shown in Figure 4.

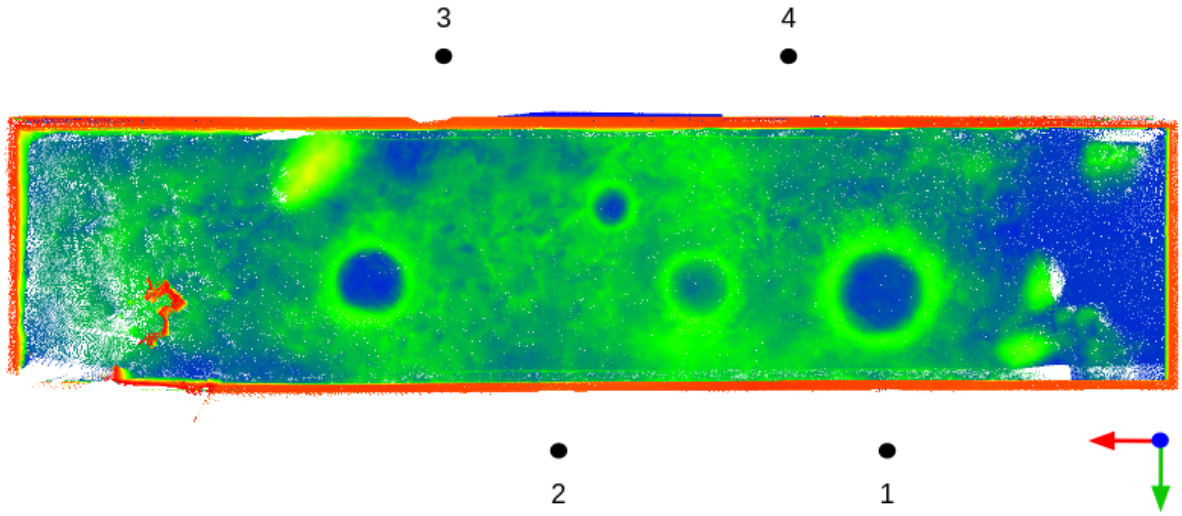
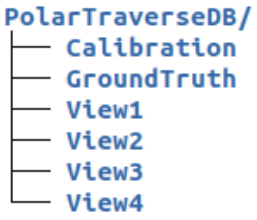


Figure 4: Processed ground truth scan of lunar lab test bed with approximate scan positions and coordinate frame (x = red, y = green, z = blue). Color indicates z height, with red indicating higher regions and blue indicating lower regions.

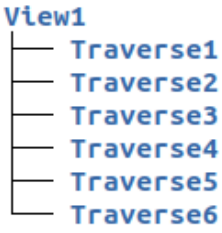
3 Dataset Organization

The dataset is organized into subfolders based on the views listed in Table 1, along with separate subfolders for the camera calibration and ground truth scans. A manifest is also provided in the top-level folder to show the parameter combinations and location IDs for each image. Each view subfolder is itself organized into subfolders based on the traverses listed in Table 2. The top-level folder structure and the traverse subfolders for an example view are shown in Figure 5.

Each traverse folder is organized into subfolders based on the distance along the test bed at which the camera jib was placed for imaging the terrain relative to the starting position for that traverse (the "00m" subfolder). Each of the position subfolders contains all of the images recorded at that position for the given view and



(a) High-level.



(b) Within one view folder.

Figure 5: Folder structure of the POLAR traverse dataset .

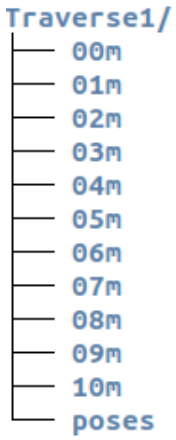


Figure 6: Folders within one traverse folder.

View	Traverse	ViewDirection	LightingAngle	CamHeight	CamPitch	DistanceAlongTestBed	ExposureTime	LocID	CamID	Path
1	1	Forward	20	1.3	14	0	1	0	L	View1/Traverse1/00m/loc0_camL_001ms.png
1	1	Forward	20	1.3	14	0	5	0	L	View1/Traverse1/00m/loc0_camL_005ms.png
1	1	Forward	20	1.3	14	0	10	0	L	View1/Traverse1/00m/loc0_camL_010ms.png
1	1	Forward	20	1.3	14	0	25	0	L	View1/Traverse1/00m/loc0_camL_025ms.png
1	1	Forward	20	1.3	14	0	50	0	L	View1/Traverse1/00m/loc0_camL_050ms.png
1	1	Forward	20	1.3	14	0	75	0	L	View1/Traverse1/00m/loc0_camL_075ms.png
1	1	Forward	20	1.3	14	0	100	0	L	View1/Traverse1/00m/loc0_camL_100ms.png
1	1	Forward	20	1.3	14	0	150	0	L	View1/Traverse1/00m/loc0_camL_150ms.png
1	1	Forward	20	1.3	14	0	200	0	L	View1/Traverse1/00m/loc0_camL_200ms.png
1	1	Forward	20	1.3	14	0	250	0	L	View1/Traverse1/00m/loc0_camL_250ms.png
1	1	Forward	20	1.3	14	0	300	0	L	View1/Traverse1/00m/loc0_camL_300ms.png
1	1	Forward	20	1.3	14	0	350	0	L	View1/Traverse1/00m/loc0_camL_350ms.png
1	1	Forward	20	1.3	14	0	400	0	L	View1/Traverse1/00m/loc0_camL_400ms.png
1	1	Forward	20	1.3	14	0	450	0	L	View1/Traverse1/00m/loc0_camL_450ms.png

Figure 7: Top portion of manifest file.

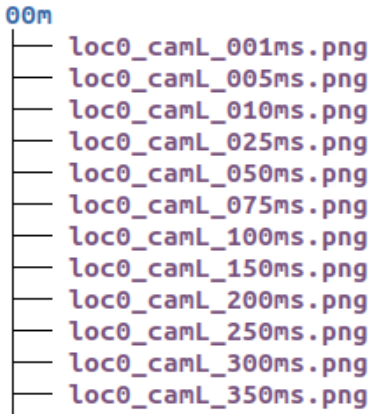


Figure 8: Example image file names from one directory of the dataset.

traverse. The camera poses for each traverse are in a separate "poses" subfolder. This structure is shown in Figure 6 for an example traverse.

The manifest file lists all images included in the dataset, along with the parameter settings used to record the image. The header and first few rows of this file are shown in Figure 7. Images can be uniquely identified by a combination of their location ID and exposure time.

4 File Formats and Details

4.1 Images

All image files included in the POLAR traverse dataset are grayscale images saved as PNG files and have a resolution of 2048 x 2048 pixels. Each image file is named "loc[LOCID]_cam[ID]_[EXP]ms.png" with [ID] taking the value of L or R to indicate the left or right camera from the stereo pair and [EXP] set to the exposure time used to take the image. [LOCID] is the location ID, which can be used to look up the associated pose in the pose files and is unique to each location.[2] Figure 8 shows some example image file names from one of the directories in the dataset.

4.2 Camera Poses

Two sets of poses are provided for each traverse within each view. The first are approximate and are based on the length along the test bed at which the images are taken, along with known parameters from the setup (camera height, pitch, and baseline). This information was used to produce files with approximate camera poses, named "approx_pose.txt" and saved in the "poses" directory for each traverse. The approximate poses are relative to the point on the ground under the first location of the left camera for a given traverse and view such that the x and y components of the first left camera pose are 0 but the z component matches the camera height. The orientations are provided as the rotation from the world frame to the camera frame, where the world frame is as shown in Figure 4. The translations are provided in the world frame's orientation.

[2]The location ID and exposure time together uniquely identify an image; location IDs are unique on view, traverse, camera (i.e. left or right), and distance along the test bed. Images taken at the same location but with different exposure times will have the same location ID. However, the left and right cameras have different location IDs since they are located at different positions in space.

The second set of poses are refined poses produced using COLMAP [3] and are saved to the same location with the names "refined_pose.txt". More information on how COLMAP was used to generate these poses is provided in Appendix A. All refined poses are computed relative to the first location of the left camera at the start of the traverse for a given view. The orientations are provided as the rotation from the world frame to the camera frame, where the world frame is as shown in Figure 4, and the translations are provided in the world frame's orientation.

Heading	Description
LOCID	Location ID for the pose
X	X translation component
Y	Y translation component
Z	Z translation component
QW	Real component of orientation quaternion
QX	X vector component of orientation quaternion
QY	Y vector component of orientation quaternion
QZ	Z vector component of orientation quaternion

Table 3: Pose file column headings and descriptions.

All pose files are formatted as tab-delimited text files with the column headings shown in Table 3. The location ID is used to uniquely determine the camera poses and can be linked directly to the image file names. Figure 9 shows an example of the first few rows of one of the pose files.

LOCID	X	Y	Z	QW	QX	QY	QZ
0	0	0	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384
1	0	-0.4	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384
2	1	0	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384
3	1	-0.4	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384
4	2	0	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384
5	2	-0.4	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384
6	3	0	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384
7	3	-0.4	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384
8	4	0	1.3	0.4353384	-0.5572077	0.5572077	-0.4353384

Figure 9: Example poses from "approx_pose.txt" file.

4.3 Camera Calibration

The camera calibrations are provided as three YAML files that describe the extrinsics, left camera intrinsics, and right camera intrinsics in formats useful for OpenCV. The first, "extrinsics.yml", contains the rotation matrix and translation vector between the two cameras, along with the average reprojection error from the calibration. The translation vector and rotation matrix are the left camera's position in the right camera's frame, and the translation vector is in meters. Figure 10 displays the content of the extrinsics calibration file.

The camera intrinsics are saved in the "left_intrinsics.yml" and "right_intrinsics.yml" files for the left and right cameras, respectively. These both contain the camera matrix and distortion coefficients for their corresponding camera. Figure 11 displays the contents of the left camera's intrinsics file.

```
%YAML:1.0
---
rotation_matrix: !!opencv-matrix
    rows: 3
    cols: 3
    dt: d
    data: [ 9.999957824489283e-01, 1.384127605741120e-04,  2.901021589618670e-03,
           -1.287498211802540e-04, 9.999944445926076e-01, -3.330796812442055e-03,
           -2.901466498043608e-03, 3.330409258625485e-03,  9.999902448855843e-01]
translation_vector: !!opencv-matrix
    rows: 3
    cols: 1
    dt: d
    data: [-0.399577424, 0.000167072, -0.000584272]
avg_reprojection_error: 0.110805
```

Figure 10: Content of the "extrinsics.yml" file.

```
%YAML:1.0
---
image_width: 2048
image_height: 2048
camera_matrix: !!opencv-matrix
    rows: 3
    cols: 3
    dt: d
    data: [1.45271e+03, 0., 0.99953e+03,
           0., 1.45288e+03, 1.03540e+03,
           0., 0., 1.]
distortion_coefficients: !!opencv-matrix
    rows: 5
    cols: 1
    dt: d
    data: [-0.016834, -0.027914, -0.000321, -0.000487, -0.001499]
```

Figure 11: Content of the "left_intrinsics.yml" file.

4.4 Ground Truth Scans

The ground truth FARO scans are provided as PLY files in ASCII format with one scalar value representing the intensity of the scan in addition to the x, y, and z position values. These scans are provided in two formats, both found in the "GroundTruth" subdirectory; the folder structure for this subdirectory is shown in Figure 12. The scan position IDs correspond to the numbers shown in Figure 4.

Raw: Four individual scans that are nearly raw, each taken at a different position relative to the test bed. Aside from converting the scans to PLY format and zeroing the minimum z value, no processing has been applied to the scans. These scans are named "position[X].ply" where X indicates the position ID (1-4) and can be found in the "GroundTruth/Raw" subdirectory.

Processed: Single composite scan generated by cropping the four individual scans to the test bed, registering each scan to one another, and downsampling to point spacing of 8 mm (the accuracy set on the FARO scanner when the scans were taken). Cropping and initial registration were performed manually, followed by fine-tuning of the registration using ICP. This single point cloud is centered on the origin and is oriented such that the positive x axis is pointing down the length of the test bed from the 0 m position of the cameras when facing in the "forward" direction (see Figure 4 for a depiction of the coordinate frame). This scan is named "registered_point_cloud.ply" and is in the "GroundTruth/Processed" subdirectory.

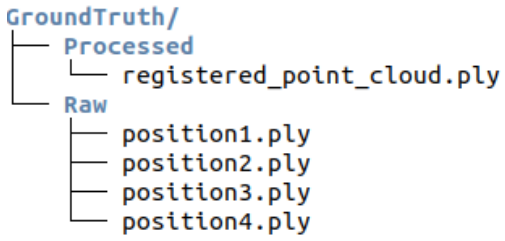


Figure 12: Folder structure for the ground truth data.

References

[1] NASA, "VIPER Mission Overview." https://www.nasa.gov/viper/overview.

[2] NASA, "Artemis III Science Definition Team Report," Tech. Rep. NASA/SP-20205009602, NASA, 2021.

[3] J. L. Schönberger and J.-M. Frahm, "Structure-from-Motion Revisited," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. Code available at https://demuc.de/colmap/.

[4] E. Labs, "Lunar highlands (lhs-1) high-fidelity moon dirt simulant." .

[5] Allied Vision, "Manta G-419." https://www.alliedvision.com/en/camera-selector/detail/manta/g-419/.

[6] Tamron, "M111FM08." https://www.tamron.vision/lenses/m111fm08/?cookie-state-change=1660854671875.

[7] Proaim, "Proaim 10' Wave-2 Jib Crane, Dolly Stand for Camera / Gimbals / Pan Tilt Head." https://www.proaim.com/products/proaim-10-wave-2-plus-telescopic-jib-arm-crane-with-cst-100-stand-portable-dolly.

[8] MATLAB, "Using the Stereo Camera Calibrator App." https://www.mathworks.com/help/vision/ug/using-the-stereo-camera-calibrator-app.html.

A Camera Pose Refinement

The refined camera positions were generated by running COLMAP [3] on a set of images taken at 50 ms exposure for each of the four views. These images covered all of the traverses and were ordered such that the camera path would trace back and forth along the test bed with different heights and pitch angles. This enabled only 4 pose refinements to be done, with post-processing to separate the poses out into individual traverses. The approximate poses were provided to COLMAP as initial poses to enable better feature matching between adjacent images.

The camera images were downsampled to half resolution (1024 x 1024) and re-calibrated to generate half-resolution calibrations that were provided to COLMAP. This step was taken to speed up processing. COLMAP's default parameter settings were applied where possible, though changes were necessary to ensure convergence of the algorithm. These changes are listed in Table 4, organized by the step in the reconstruction process in which the parameter is used.

Feature matching	
Method	Spatial
Maximum number of neighbors	15
Maximum search distance	5 m
GPS	No
Ignoring z	No
Maximum error	3
Minimum inlier ratio	0.4
Minimum number of inliers	50
Sparse reconstruction	
Initial maximum forward motion	1 m
Initial minimum triangle degrees	5°
Maximum pose error	6
Minimum inliers for pose	100
Minimum inlier ratio for pose	0.6
Ignoring two-view tracks	No
Minimum angle for triangulation	2°
Refining camera parameters	Yes

Table 4: Parameter values used to produce poses for the scenes included in this dataset from COLMAP. Only parameters that were changed from their default value are included. Headings indicate the processing step in which a parameter is found.