ASRL-2012-ABL001
Rev: 1.0
September 25, 2013

The Gravel Pit Lidar-Intensity Imagery Dataset

Sean Anderson*, Colin McManus†, Hang Dong*, Erik Beerepoot* and Timothy D. Barfoot*

*University of Toronto Institute for Aerospace Studies, †University of Oxford

{sean.anderson,erik.beerepoot}@mail.utoronto.ca

{hang.dong,tim.barfoot}@utoronto.ca, colin@robots.ox.ac.uk

Abstract

This dataset contains intensity and range data collected using a high-framerate, two-axis scanning lidar over ten individual traversals of the same 1.1km path. The experiment was conducted over a full diurnal cycle at a planetary analogue in Sudbury, Ontario, Canada and should be of interest to researchers who develop algorithms for visual odometry, simultaneous localization and mapping (SLAM) or place recognition in three-dimensional, unstructured, and natural environments. Catering to the strength of state-of-the-art SLAM techniques, this dataset creates ample opportunity for loop closure; in addition to having multiple traversals of the same path, the trajectory was specifically chosen to include both small- and large-scale loops. The lidar scans were taken with a 480×360 resolution at 2Hz, while driving roughly 0.3-0.4 meters per second; therefore, one of the challenges in using this dataset is to compensate for the motion distortion present in the data (resulting from the 'rolling-shutter' effect). Ground truth position is provided by means of a Thales DG-16 Differential GPS unit.

1 Introduction

The Gravel Pit Lidar Intensity Imagery Dataset is a collection of 77,754 high-framerate laser range and intensity images gathered at a planetary analogue environment in Sudbury, Ontario, Canada, as seen in Figure 1. The data were collected during a visual teach and repeat experiment (McManus et al., 2012) in which a 1.1km route was taught and then autonomously retraversed (i.e., the robot drove in its own tracks) every 2-3 hours for 25 hours. The dataset is subdivided into the individual 1.1km traversals of the same route, at varying times of day (ranging from full sunlight to full darkness).

Figure 1: The ROC6 at the Ethier Sand and Gravel pit in Sudbury, Ontario, Canada.

The lidar scanner used in this dataset, the Autonosys LVC0702, captured both intensity and range images with a resolution of 480x360 at 2Hz. The unique output of this sensor makes this dataset an interesting candidate for 6D motion estimation in unstructured environments. Not only can the range data be used for scan alignment algorithms, but appearance-based features extracted from the intensity imagery have enabled visual estimation schemes to also be possible with this sensor (McManus et al., 2011). To encourage the use of this dataset and lower the barrier to entry, we provide both the raw features outputted by our Speeded-Up Robust Feature (SURF) implementation and a set of temporally tracked features that can be used for motion estimation.

This dataset should be useful for field robotics researchers developing algorithms for visual odometry, simultaneous localization and mapping (SLAM) or place recognition in three-dimensional, unstructured, natural terrain. Unlike many of the prominent 3D laser scan datasets, such as the Osnabrück Robotic 3D Scan Repository (Nüchter and Lingemann, 2009) and The Canadian Planetary Emulation Terrain 3D Mapping Dataset (Tong et al., 2012), which are suited for survey-style mapping and take long panoramic scans from a select number of static locations, this dataset is geared towards using high-rate scanning lidar as an active localization method. Therefore, regardless of whether the data is being used in a scan-alignment or sparse-visual-feature method, such as bundle adjustment, one of the greatest challenges in using this dataset will be compensating for the motion distortion present in the scans, resulting from the 'rolling-shutter' effect. This distortion is due to vehicle motion and the scanning nature of lidar, akin to a slow rolling shutter camera.

A similar style of online 3D laser scan acquisition is provided by the Velodyne-3D lidar scanner in the Ford Campus Vision and Lidar Data Set (Pandey et al., 2011); however, the dataset was acquired in an urban area, and the vertical resolution of the Velodyne-3D scanner is unsuitable for use with our intensity-based feature extraction scheme. To the authors' knowledge, the only other dataset to provide lidar intensity information is the New College Vision and Laser Data Set (Smith et al., 2009); however, due to the static vertical mounting of the laser scanners, the laser data alone cannot be used for odometric estimation.

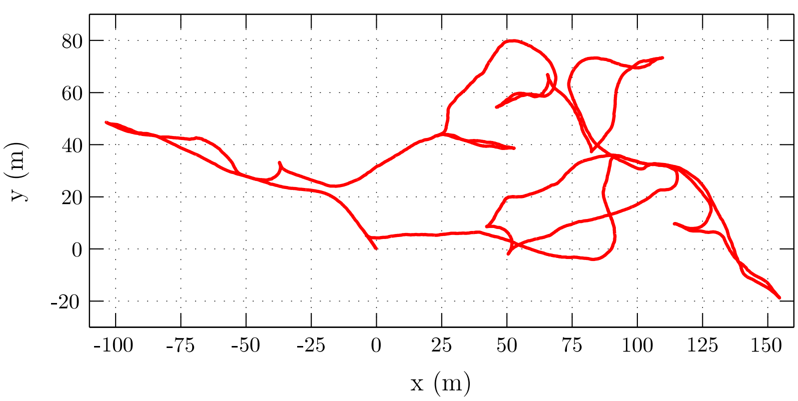


Figure 2: GPS during a single 1.1km path traversal.

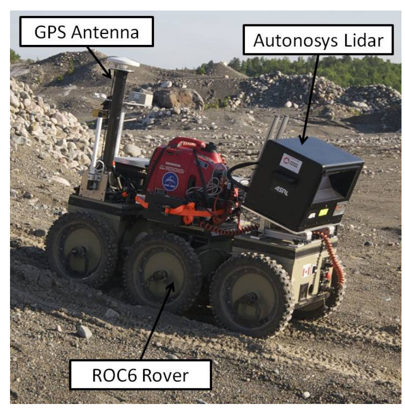


Figure 3: The ROC6 mobile platform

Another key feature of this dataset is the opportunity for place recognition and loop closure. As seen in Figure 2, the shape of the path traversed in this dataset provides both small local loop closures, as well as a large-scale loop closure over the whole 1.1km traverse. Furthermore, we provide 10 runs of the same 1.1km route (with the robot driving in its own tracks), allowing for both place recognition and loop closure between runs.

The datasets, detailed packaging descriptions, and videos are available at *http://asrl.utias.utoronto.ca/datasets/abl-sudbury/*. The various data products are provided as either human-readable text files or images and are accompanied by Matlab parsing scripts for ease of use.

2 Hardware Setup

The ROC6, seen in Figure 3, is an articulated and skid-steered mobile platform consisting of three individual pods; the front and rear pods are able to pitch and roll relative to the central one. During each traversal of the 1.1km route, the rover travelled at roughly 0.3-0.4 meters per second.

The payloads most relevant to this dataset are the high-framerate Autonosys LVC0702 lidar, and a Thales DG-16 Differential GPS unit. The Autonosys LVC0702 lidar provides 500,000 measurements per second with a 15-bit intensity at a maximum range of 53.5m. In this dataset, the lidar was configured to have a 90°H/30°V field of view, and capture images with a resolution of 480×360 at 2Hz. The Thales DG-16 Differential GPS unit has a Circular Error Probability (CEP) of 0.4m, with 95% of measurements occuring within 0.9m.

3 Overview of Datasets

Each dataset in this collection corresponds to a unique traversal of the same 1.1km route (conducted at different times of day). For ease of use, the data have been post-processed and packaged into a few different products. This section will provide an overview of the available dataproducts. The specifics of each traversal can be found in Table 1.

The first major dataproduct we provide is the sequence of Autonosys imagery, generated from the raw sensor data, in the Tagged Image File Format (TIFF). Each intensity image in the sequence is accompanied by a corresponding azimuth, elevation, range, mask and timestamp image. TIFF was chosen as it supports 32/64-bit floating point images and is simple to load using either Matlab, or OpenCV (which leverages LibTiff).

The second major dataproduct we provide is SURF features and frame-to-frame matches for all the intensity imagery. SURF features are extracted using a GPU-accelerated SURF implementation and two sets of frame-to-frame matches are provided. The first set of matches is simply the initial guesses based on only the SURF descriptor. The second set of matches is the inliers after a being passed through a RANdom SAmple Consensus (RANSAC) algorithm that accounts for the motion distortion in the image.

4 Description of Data Products

In this section, we detail the format of the data in addition to specifics such as experimental considerations and post-processing details. Each dataset contains a set of folders corresponding to the various data products. Each of these folders contain a full sequence of either TIFF images or comma-delimited, human-readable text files. All comma-delimited text files begin with a single comma-delimited header line that contains titles corresponding to the data items.

4.1 Dataset Header File

Each dataset contains a single human-readable, comma-delimited header file with information pertaining to the contents of the dataset. In order, each line of the header file contains:

- **id**: The identification number used throughout the dataset to associate data belonging to a specific image stack.
- **timestamp**: Although each image is captured over a period of time, this is nominal time we consider the frame to be captured at. It is calculated as the average of the first and last measurement timestamp. This number is expressed in seconds since the beginning of the experiment.
- **validpx**: Due to packet loss, some images may have a few blank pixel rows. This value is the fraction of valid data in the frame (floating-point between 0 and 1).

Traverse	Number of Frames	Start Time	End Time	Special Notes
Teach 1	6880	19:45:xx	xx:xx:xx	All other runs track this *teach* traversal.
Run 1	7039	23:03:27	00:03:39	none
Run 4	6741	05:00:28	05:56:26	Missing approximately 250 meters of DGPS data at the beginning of the traversal.
Run 5	8679	09:47:12	10:57:13	Missing approximately 100 meters of DGPS data at the beginning of the traversal.
Run 6	9694	11:51:36	13:20:19	After frame 5879, there is a break in the imagery where the robot was paused and the lidar scanner was reset. The translation between the break is only 1.5-2 meters and significant overlap between the frames still exists.
Run 7	9644	14:15:54	15:35:51	none
Run 8	8691	16:25:05	17:32:41	Missing approximately 50 meters of DGPS data at the beginning of the traversal.
Run 9	6456	18:24:19	19:18:41	none
Run 10	7863	20:31:06	21:37:36	none
Run 11	6067	22:58:43	23:50:06	none

Table 1: This table gives an overview of the available traverse data. Note that runs 2 and 3 have been excluded from the dataset as they were compromised by hardware malfunction and data loss. Preview videos for each traverse, including feature tracks, are available on the website.

4.2 GPS Data File

The GPS data file is a human-readable, comma-delimited text file containing the GPS coordinate at each frame capture. During the collection phase, the GPS and lidar were not synchronized. To account for this, the GPS coordinates that are provided have been interpolated to occur at the nominal timestamp at each frame. Following the header line, each line of the GPS file contains:

- **id**: The frame identification number.
- **x, y, z**: Recentered UTM coordinates (in meters).

4.3 Alignment Matrices

As depicted in Figure 4, the three frames related to the measurement data are the sensor frame, $\underrightarrow{\mathcal{F}}_c$, the GPS frame, $\underrightarrow{\mathcal{F}}_{\text{gps}}$, and the inertial frame, $\underrightarrow{\mathcal{F}}_i$. This dataset uses homogeneous transformation matrices to express the translation and rotation between frames. For example, a point in $\underrightarrow{\mathcal{F}}_b$ can be transformed into $\underrightarrow{\mathcal{F}}_a$ using the matrix $\mathbf{T}_{a,b}$ in the following manner:

$$\begin{bmatrix}\mathbf{p}_a^{l,a}\\ 1\end{bmatrix} = \mathbf{T}_{a,b}\begin{bmatrix}\mathbf{p}_b^{l,b}\\ 1\end{bmatrix} = \begin{bmatrix}\mathbf{C}_{a,b} & \boldsymbol{\rho}_a^{b,a}\\ \mathbf{0}^T & 1\end{bmatrix}\begin{bmatrix}\mathbf{p}_b^{l,b}\\ 1\end{bmatrix}, \quad \text{or using components,} \quad \mathbf{p}_a^{l,a} = \mathbf{C}_{a,b}\mathbf{p}_b^{l,b} + \boldsymbol{\rho}_a^{b,a},$$

where $\mathbf{p}_a^{l,a}$ is the vector from $\underrightarrow{\mathcal{F}}_a$ to point l, expressed in $\underrightarrow{\mathcal{F}}_a$, similarly $\mathbf{p}_b^{l,b}$ is the vector from $\underrightarrow{\mathcal{F}}_b$ to point l, expressed in $\underrightarrow{\mathcal{F}}_b$, $\mathbf{C}_{a,b}$ is the rotation matrix from $\underrightarrow{\mathcal{F}}_b$ to $\underrightarrow{\mathcal{F}}_a$, and $\boldsymbol{\rho}_a^{b,a}$ is the translation from $\underrightarrow{\mathcal{F}}_a$ to $\underrightarrow{\mathcal{F}}_b$, expressed in $\underrightarrow{\mathcal{F}}_a$. More detailed and practical examples of using homogeneous transformation matrices can be found in the example estimation Matlab code.

The first matrix we provide is the 4×4 homogeneous transformation matrix relating the sensor frame and GPS frame, $\mathbf{T}_{c,\text{gps}}$. For simplicity and due to the scale of the CEP, this transform is assumed static, and provided only for the nominal

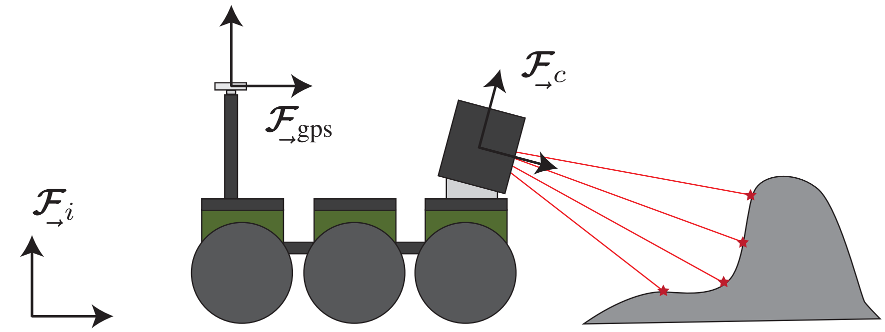


Figure 4: Coordinate Frame Definitions

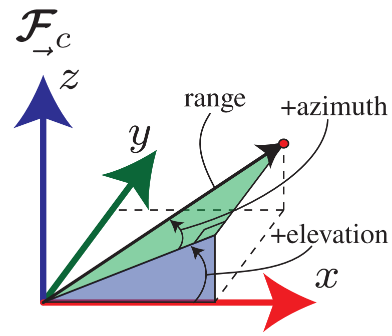
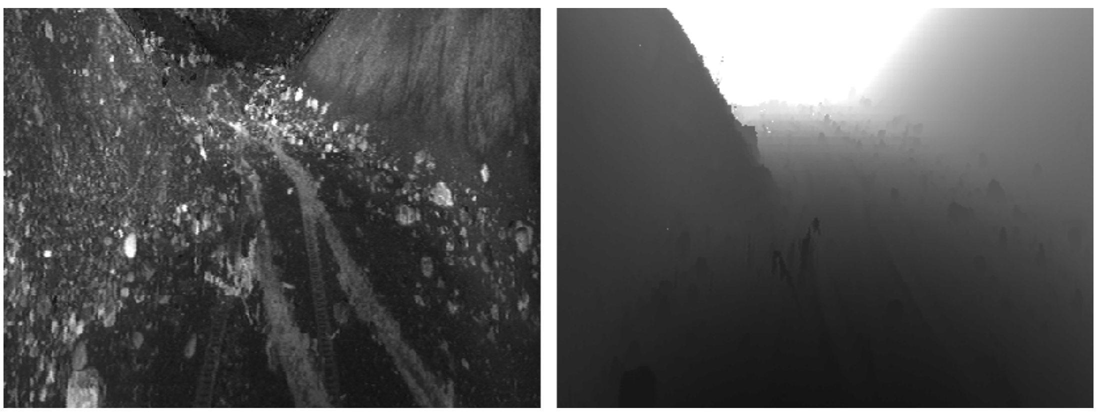


Figure 5: Camera model



(a) Example 8-bit intensity image

(b) Example 32-bit range image

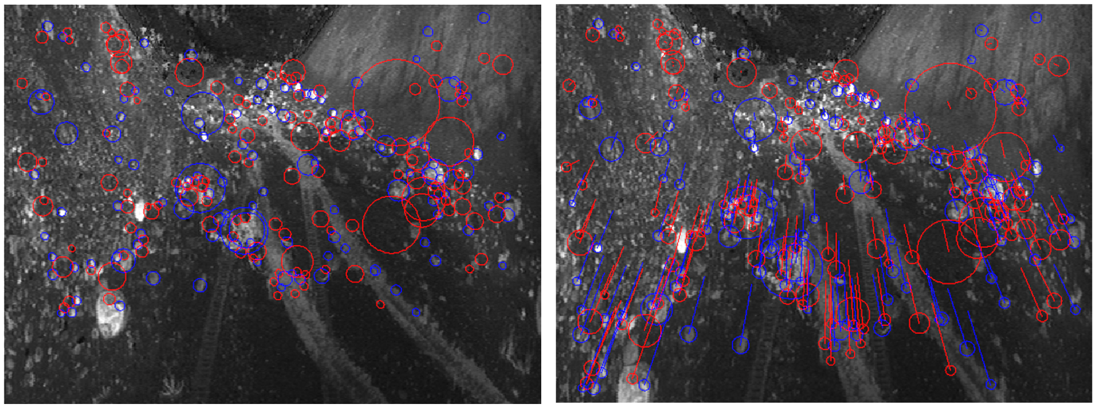
Figure 6: A typical pair of Autonosys intensity and range images

position of the pods. Second, each traverse dataset contains a transformation matrix, $\mathbf{T}_{\text{gps},i}$, to bring the initial local GPS frame into angular alignment with the inertial GPS data. This alignment matrix is calculated by performing a simple point-to-point, least-squares optimization between the first 30 meters of our visual odometry estimate and the GPS data.

The file format for matrices, `matrix_<name>.txt`, is straightforward. The first line contains the comma separated number of rows and columns in the matrix. The following lines contain the floating-point data of the matrix (comma separated for columns and and line separated for rows).

4.4 Image Stacks

For each frame listed in the header file, there exists a set of `.tif` images that make up a single *image stack*. Each image stack has a resolution of 480×360 and consists of two 8-bit unsigned integer images (post-processed intensity and mask), one 16-bit unsigned integer image (raw intensity), three 32-bit floating point images (azimuth, elevation and range) and one 64-bit floating point image (time). An example intensity and range image can be seen in Figure 6. The 8-bit intensity image is a range corrected version of the 16-bit raw intensity image. Due to the occurance of packet loss, the mask image has been provided to mark valid pixel data; 255 corresponding to valid data, and 0 to invalid data. The raw azimuth, elevation and range images make up the geometric portion of the scans and the associated spherical camera model is depicted in Figure 5. Note that these raw measurements do not yet include the intrinsic calibration, which was performed using the generalized distortion model found in (Dong et al., 2013). The undistortion function is straightforward and has been made available in the Matlab code. Finally,



(a) Intensity image k with SURF features **(b)** Image $k + 1$ with filtered matches to image k

Figure 7: A pair of sequenced Autonosys intensity images with SURF features and tracks.

the time image provides per-pixel timing information for the measurements (in seconds since the beginning of the experiment).

4.5 SURF Feature File

The SURF feature files contain a list of SURF features extracted from the Autonosys lidar intensity data. There is one SURF feature file for every image stack in the dataset. When using the sub-pixel (u, v) coordinate to extract measurements from the image stacks, the four surrounding pixels were used for bilinear interpolation. Additionally, the azimuth, elevation, and range measurements have already been idealized using the supplied intrinsic calibration model (see the `autonosys_apply_calib` helper function). An example intensity image with SURF features can be seen in Figure 7a.

Within each comma-delimited SURF feature file, we have recorded, in order, the horizontal and vertical pixel coordinate, 8-bit intensity, azimuth, elevation, range, time, feature size, feature response strength, feature orientation, octave, angular response strength, covariances, laplacian and the 64 floating-point value SURF descriptor. Additional information about each of these values can be found on the website.

4.6 SURF Feature Match File

The feature match file contains a list of indices that relate SURF features in sequential frames. Two types of match files have been provided. The first are the raw matches, which are based solely on the SURF feature descriptors. The second are the filtered matches, which provide only inlier matches based on a RANSAC algorithm that considers motion distortion. An example intensity image with filtered SURF feature tracks can be seen in Figure 7b. Each feature match file contains a comma-delimited list of index pairings that specify the matches between frames k and $k + 1$.

5 Helpful Tools

This dataset is accompanied by a set of useful Matlab scripts aimed at reducing the amount of effort required to start using this data. These scripts include: parsing code for all comma-delimited files, a loading function for image stacks, display functions, a function to apply intrinsic calibration, conversions between spherical and Cartesian coordinates and three pieces of example code. The first example opens and displays image stacks, the second displays SURF features and tracks, the third is posed toward setting up an estimation problem between sequential images. Refer to the website for further details.

Acknowledgements

The collection of this data would not have been possible without the support of many people. In particular, we would like the thank the staff of the Ethier Sand and Gravel in Sudbury, Ontario, Canada for allowing us to conduct our field tests on their grounds and Dr. James O'Neill from Autonosys for his help in preparing the lidar sensor for our field tests. We also wish to thank DRDC Suffield, MDA Space Missions, the NSERC, the Canada Foundation for Innovation, and the Canadian Space Agency for providing us with the financial and in-kind support necessary to conduct this research.

References

Dong, H., Anderson, S., and Barfoot, T. D., "Two-axis Scanning Lidar Geometric Calibration using Intensity Imagery and Distortion Mapping," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Karlsruhe, Germany, 2013.

McManus, C., Furgale, P. T., and Barfoot, T. D., "Towards Appearance-based Methods for Lidar Sensors," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 1930–1935, Shanghai, China, 2011.

McManus, C., Furgale, P. T., Stenning, B. E., and Barfoot, T. D., "Visual Teach and Repeat Using Appearance-Based Lidar," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 389–396, St. Paul, USA, 2012.

Nüchter, A. and Lingemann, K., "Robotic 3D Scan Repository," http://kos.informatik.uni-osnabrueck.de/3Dscans/, 2009.

Pandey, G., McBride, J., and Eustice, R., "Ford Campus Vision and Lidar Data Set," *International Journal of Robotics Research*, 30(13):1543–1552, 2011.

Smith, M., Baldwin, I., Churchill, W., Paul, R., and Newman, P., "The New College Vision and Laser Data Set," *International Journal of Robotics Research*, 28(5):595–599, 2009.

Tong, C. H., Gingras, D., Larose, K., Barfoot, T. D., and Dupuis, E., "The Canadian Planetary Emulation Terrain 3D Mapping Dataset," *International Journal of Robotics Research*, 2012.