

# Towards Appearance-Based Methods for Lidar Sensors

Colin McManus, Paul Furgale, and Timothy D. Barfoot

**Abstract**—Cameras have emerged as the dominant sensor modality for localization and mapping in three-dimensional, unstructured terrain, largely due to the success of sparse, appearance-based techniques, such as visual odometry. However, the Achilles' heel for all camera-based systems is their dependence on consistent ambient lighting, which poses a serious problem in outdoor environments that lack adequate or consistent light, such as the Moon. Actively illuminated sensors on the other hand, such as a light detection and ranging (lidar) device, use their own light source to illuminate the scene, making them a favourable alternative in light-denied environments. The purpose of this paper is to demonstrate that the largely successful appearance-based methods traditionally used with cameras can be applied to laser-based sensors, such as a lidar. We present two experiments that are vital to understanding and enabling appearance-based methods for lidar sensors. In the first experiment, we explore the stability of a representative keypoint detection and description algorithm on both camera images and lidar intensity images collected over a 24 hour period. In the second experiment, we validate our approach by implementing visual odometry based on sparse bundle adjustment on a sequence of lidar intensity images.

## I. INTRODUCTION

Cameras have emerged as the dominant sensor modality for localization and mapping in three-dimensional, unstructured terrain, largely due to the success of sparse, appearance-based techniques. Local appearance-based approaches search for a sparse set of descriptive regions in an image and describe these regions based on the intensity patterns of the surrounding area in image space (e.g., Scale Invariant Feature Transforms (SIFT) [1]). A unique fingerprint is then assigned to each of these descriptive regions, called *keypoints*, so they can be identified and matched in subsequent images. Sparse appearance-based techniques have been combined with fast optimization methods to build a number of fundamental autonomy-enabling algorithms such as visual odometry (VO) [2], visual teach and repeat (VT&R) [3], and visual simultaneous localization and mapping (VSLAM) [4].

However, all camera-based systems are inherently sensitive to the amount of ambient light, which poses a serious problem in outdoor environments that lack adequate and consistent light, such as the Moon. Even on Earth, changes in ambient lighting over the course of a day can result in failures to recognize pre-traversed areas [3]. Thus, it would be attractive to move to an actively illuminated sensor to reduce external lighting dependencies. Our aim is to take the lessons learned from the success of camera-based systems and apply the same methods to systems that use laser-based

The authors are members of the Autonomous Space Robotics Lab at the Institute for Aerospace Studies, University of Toronto, 4925 Dufferin Street, Toronto, Ontario, Canada { colin.mcmanus, paul.furgale, tim.barfoot }@utoronto.ca

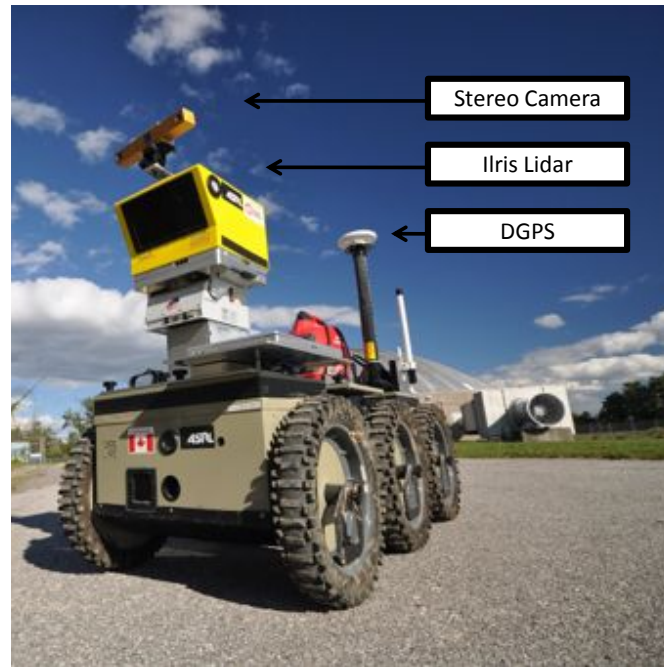


Fig. 1. Our field robot with a Bumblebee XB3 stereo camera, a Thales DG-16 Differential GPS unit and an Optech Iris 3D lidar sensor, which is a three-dimensional surveying lidar that has a vertical and horizontal field of view of 40°.

sensors, such as the lidar depicted in Figure 1. The key insight is that, in addition to range data, lidar sensors also provide intensity information, allowing one to construct an intensity image that looks very similar to a standard grayscale camera image.

Recognizing that laser intensity images<sup>1</sup> provide a grayscale image of a scene is not a new idea. Kretschmer et al. [5] point out that in surveying, the intensity images are often used by the surveyor to obtain a photo-realistic impression of the scanned area. In fact, most commercial surveyors use various reflective markers in the scene to act as tie points between different scan positions [6]. In an effort to automate this scan registration process, Bohm et al. [7] developed a marker-free method for point cloud registration that uses point correspondences from the intensity images to estimate the rigid body transformations. SIFT features were extracted from the intensity images and Random Sample and Consensus (RANSAC) [8] was used for outlier detection. In order to dampen the areas of low and very high reflectance, histogram equalization was used on all of the raw intensity images [7].

<sup>1</sup>Also referred to as *reflectance images*.

In the mobile robotics literature, few have actually incorporated intensity information from a laser sensor for motion estimation. Neira et al. [9] developed a sensor fusion technique in planar environments using their variant of the Extended Kalman Filter (EKF), called the SPfilter, which incorporated both range and intensity data to localize against a known map. Guivant et al. [10] described a SLAM system that used the intensity data from their lidar to identify reflective markers on landmarks in the environment, which simplified the data association problem.

The most relevant research to date comes from May et al. [11], who developed 3D mapping and egomotion estimation techniques using a Swiss Ranger time-of-flight (ToF) camera, which measures distances based on the phase-shift principle. Unlike laser scanners, such as the lidar used in our study, the Swiss Ranger uses an array of 24 LEDs to simultaneously illuminate a scene, offering the advantage of higher framerates and an easier camera model for image formation. However, ToF cameras often have a limited field of view, short maximum range, and are very sensitive to environmental noise [11].

Using intensity images generated from their ToF camera, May et al. employed two feature-based methods for motion estimation: a KLT-tracker and frame-to-frame VO using SIFT features. Their results indicated that the SIFT approach yielded more accurate motion estimates than the KLT approach, but less accurate than their ICP method, which used a network-based global relaxation algorithm [12].

Although May et al. demonstrated that frame-to-frame VO might be possible with a ToF camera, the largest environment in which they tested was a 20m long indoor hallway, with no groundtruth. Thus, a number of important questions still remain. In particular, we set out to answer the following questions:

- 1) How stable are descriptive features in lidar intensity images under changes in ambient light?
- 2) Can we perform VO using a scanning laser rangefinder in a stop-scan-go methodology with comparable results to stereo VO?

As this paper will show, it is indeed possible to use the intensity images from a lidar sensor and apply appearance-based techniques that have been traditionally used with camera imagery.

This paper is organized as follows. In section II we describe the lidar sensor used in this study and discuss the image formation and image processing methods required to convert raw lidar intensity data into grayscale intensity images. In section III, we explore the stability of a representative keypoint detection and description algorithm on both camera images and lidar intensity images collected over a 24 hour period outdoors. In section IV we validate our approach by implementing visual odometry based on sparse bundle adjustment on a sequence of lidar intensity images. Stereo VO results are provided for comparison and DGPS was used for groundtruth. In section V we present a discussion of our findings, with particular emphasis on the limitations and the current obstacles encountered with lidar sensors.

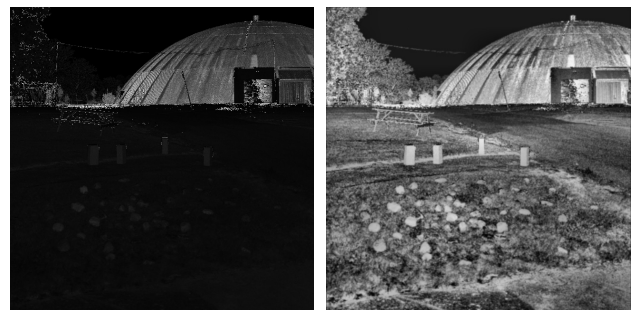
## II. PRELIMINARIES

### A. Hardware

The lidar sensor used in this study was an Optech Ilris 3D, which has a maximum horizontal and vertical field of view (FOV) of 40 degrees. The Ilris is a ToF lidar that can acquire 2,500 points/second and offers a large dynamic scanning range from 3m to over 1km, with sub-centimeter accuracy up to 100m (for our experiments, we restricted the maximum range to 500m). Since it is a surveying sensor, the Ilris was not designed for high framerate data acquisition and consequently, we had to run our robot in a stop-scan-go fashion for our localization experiment. Having said this, there are currently several options available for high framerate 3D lidar sensors, such as the LVC0702 lidar from Autononsys<sup>2</sup>, and the techniques described in this paper are applicable to these sensors as well. In addition to the lidar, we also used a Point Grey Research Bumblebee XB3 stereo camera for our visual odometry experiment, in order to provide a comparison with our lidar VO. For groundtruth, we used a Thales DG-16 Differential GPS unit.

### B. Image formation

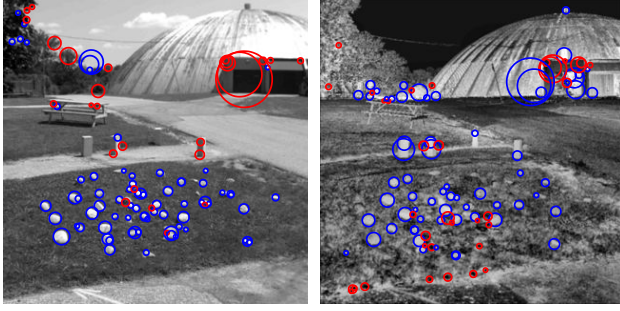
To form images from the raw lidar data, we need to develop a camera model. As this lidar provides equally spaced samples in azimuth and elevation, we were able to use the raw data directly in a spherical camera model (see Figure 2(a) for an example of a raw lidar intensity image). Once the raw intensity image is formed, image processing is required to equalize the areas of high and low reflectance. Taking a similar approach as Bohm et al. [7], we use adaptive histogram equalization and then smooth the image with a Gaussian low-pass filter. The processed image is shown in Figure 2(b). The justification for this image processing is described in the next section, since our decision was based on the results from our light sensitivity analysis.



(a) Raw lidar intensity image. The metal dome in the background is the most prominent object in the image, equalization and a Gaussian low-pass filter. The adaptive histogram equalization dampens the areas of high and low reflectance, bringing into focus many rich features in the foreground.

Fig. 2. Illustrating the image processing required to transform the raw intensity image into a textured grayscale image. All images are 640x640.

<sup>2</sup><http://www.autonosys.com/lidar.htm>



(a) Camera intensity image with the 100 strongest keypoints plotted. Obviously, in the case of the camera images, the number of keypoint matches drops off significantly as the ambient light changes. (b) Processed lidar intensity image, with the 100 strongest keypoints plotted. Although the intensity images did change as the ambient light changed, the imagery was relatively stable over a 24h period.

Fig. 3. Camera and lidar intensity images of the same scene, taken at the same time of day with sample SURF keypoints shown. Blue circles represent light-on-dark blobs and red circles represent dark-on-light blobs.

### C. Keypoint detection

We used a GPU implementation of the SURF [13] algorithm<sup>3</sup> to find keypoints in every camera and lidar intensity image. The SURF algorithm was used due to its success in previous work [3]. Figure 3 shows an example of the strongest 100 SURF keypoints detected in a camera/lidar intensity image.

## III. LIGHT SENSITIVITY ANALYSIS

In this section, we describe an experiment to analyze the stability of sparse appearance-based methods for lidar intensity images and compare them to camera intensity images, under a variety of lighting conditions. We begin with a description of the experiment and conclude with our results.

### A. Experiment description

This experiment was conducted outdoors over a 24 hour period and consisted of taking camera images and lidar scans at half-hour intervals of an outdoor scene. Since the purpose of this experiment was to analyze the impact light changes had on lidar/camera intensity images, the lidar remained stationary for the entire experiment.

To quantify the stability of the algorithm over the 24 hours, we performed exhaustive matching of every image to every other image in the dataset. Given  $N_{i,j}$  matches between image  $i$  and image  $j$ , the similarity score,  $S_{i,j}$ , is

$$S_{i,j} := \frac{N_{i,j}}{N_{\max}}, \quad (1)$$

where  $N_{\max}$  is the maximum number of possible matches (500 in this case). Keypoints matches were identified if their position in image space and descriptor space was the same to within some tolerance and three-point RANSAC [8] was used for outlier rejection.

<sup>3</sup><http://asrl.utias.utoronto.ca/code/gpusurf/index.html>

TABLE I

MEAN SIMILARITY SCORES FOR CAMERA/LIDAR INTENSITY IMAGES WITH DIFFERENT IMAGE PROCESSING METHODS.

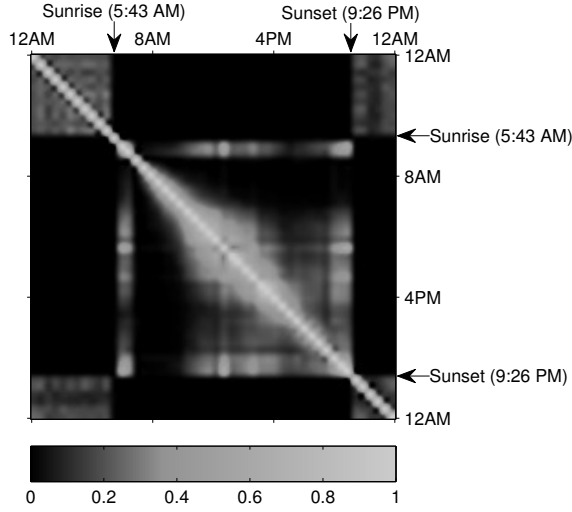
Image Processing	Camera	Lidar
(i) None	0.092±0.150	0.329±0.101
(ii) Histogram Eq.	0.069±0.135	0.443±0.127
(iii) Gaussian Filter	0.094±0.151	0.375±0.110
(iv) Histogram Eq. + Gaussian Filter	0.073±0.150	0.502±0.161

### B. Results

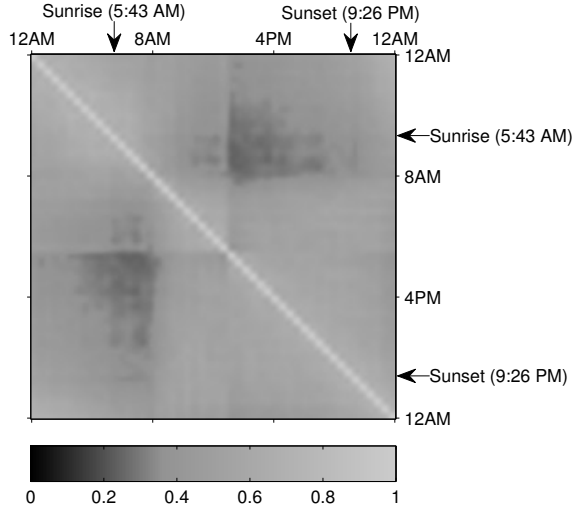
We have created similarity matrices composed of the similarity scores between all pairs of images. Figure 4 shows the similarity matrices for the camera and lidar over the entire 24 hour period, where the light values represent a greater number of matches and the dark values represent fewer matches. We compared camera and lidar similarity matrices for four different image processing options: (i) no image processing, (ii) adaptive histogram equalization, (iii) Gaussian low-pass filter, and (iv) adaptive histogram equalization and a Gaussian low-pass filter (only the best two similarity matrices are included as figures). Of the four options, we determined that applying adaptive histogram equalization and a Gaussian low-pass filter was the best for the lidar intensity images, while applying just the Gaussian low-pass filter provided a slight improvement with the camera images. Table I shows the mean similarity scores and standard deviation for all image processing options.

As expected, for the camera, the number of keypoint matches between daytime and nighttime images drops off significantly (in fact, goes to zero), which produces the dark boxes in the similarity matrix. Referring to the camera's similarity matrix, it is evident that in general, the similarity score for the camera images is quite low, illustrating the camera's sensitivity to ambient light. This means that methods such as visual teach and repeat would show a decrease in performance under changing lighting conditions. The similarity score ranged from as low as zero to as high as 0.810 with a mean score of 0.094±0.151.

In the case of the lidar intensity images, we found that the similarity score does drop off from day to night, indicating a potential sensitivity to ambient light. In fact, the similarity scores exhibited a larger deviation than expected, ranging from as low as 0.250 to as high as 0.738 throughout the entire 24 hour period (the mean score was 0.502±0.161). However, even for the most drastic changes in ambient light (i.e., from light to dark), we were still able to find at least 125 keypoint matches, which we found to be more than sufficient for motion estimation (this claim is substantiated in the next section, with a plot that shows the number of inlying keypoints over a 200m traverse). It should be noted that there were dynamic environmental factors that contributed to the drop off in the number of keypoint matches from daytime to nighttime, such as tree movement due to the wind and heliotropism. These phenomena can be seen in the time-lapsed videos provided with this paper.



(a) Camera intensity image similarity matrix. The similarity score ranged from 0-0.810 (i.e., 0-405 keypoint matches) with a mean score of  $0.094 \pm 0.151$ . Note the interesting structure in the similarity matrix, where there are sharp transitions from a non-zero score to a score of zero due to sunrise and sunset.



(b) Lidar intensity image similarity matrix. The similarity score ranged from 0.250-0.738 (i.e., 125-369 keypoint matches) with a mean score of  $0.502 \pm 0.161$ . Note that the lowest similarity scores occur between scans separated by approximately 12 hours.

Fig. 4. Similarity matrices over a 24 hour period. The similarity score is  $S_{i,j} := N_{i,j}/N_{\max}$ , where  $N_{i,j}$  is the number of matches between image  $i$  and image  $j$  and  $N_{\max}$  is the maximum number of possible matches, which was 500 in our experiment.

#### IV. VISUAL ODOMETRY

As a proof of concept of the efficacy of our approach, we implemented visual odometry using SURF keypoints tracked through an intensity image sequence. Visual odometry based on local bundle adjustment is a core algorithm in the emerging VSLAM paradigm [4], [14].

We have implemented the sliding window VO algorithm described in [15], and use the code to produce two motion estimates: one from stereo camera data, and the other from lidar scans. Processing the lidar data only requires differences in two areas: (i) keypoint formation, and (ii) error terms

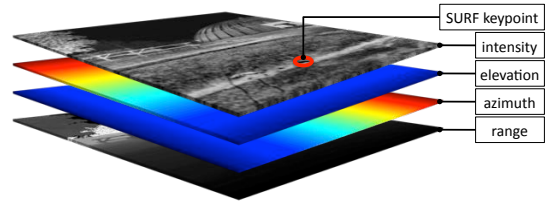


Fig. 5. The image stack generated from the raw lidar data. SURF keypoints are found in image space at sub-pixel locations and bilinear interpolation is used to find the azimuth, elevation, and range of the keypoint. Linearized error propagation from image space to azimuth/elevation/range is then used to determine the uncertainty of the measurement.

and associated Jacobians. The rest of the code blocks—keypoint tracking, hypothesis generation and outlier rejection using RANSAC, and sparse bundle adjustment—are identical to the traditional camera-based approach. Hence, we will restrict ourselves to describing the measurement formation and error terms.

##### A. Measurement and Error Terms from Lidar Data

The output of image formation (Section II-B) is a stack of images,  $\mathcal{I}$ —intensity ( $\mathcal{I}_\ell$ ), azimuth ( $\mathcal{I}_\theta$ ), elevation ( $\mathcal{I}_\phi$ ), and range ( $\mathcal{I}_r$ )—derived from the raw lidar output and shown in Figure 5. The stack may be evaluated at any integer row,  $r$ , and column,  $c$ , as,  $\mathcal{I}_{rc}$ , a  $4 \times 1$  column,

$$\mathcal{I}_{rc} := \mathcal{I}(r, c) = [\ell_{rc} \ \theta_{rc} \ \phi_{rc} \ r_{rc}]^T, \quad (2)$$

where  $\ell_{rc}$ ,  $\theta_{rc}$ ,  $\phi_{rc}$ , and  $r_{rc}$  are the scalar intensity, azimuth, elevation, and range stored at this location in the image stack. Although intensity and range are correlated, we make the simplifying assumption that the elements of each image are independent, identically-distributed samples such that

$$\mathcal{I}_{rc} = \bar{\mathcal{I}}_{rc} + \delta\mathcal{I}_{rc}, \quad \delta\mathcal{I}_{rc} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \quad (3a)$$

$$\mathbf{R} := \text{diag} \{ \sigma_\ell^2, \sigma_\theta^2, \sigma_\phi^2, \sigma_r^2 \}, \quad (3b)$$

where  $\bar{\mathcal{I}}_{rc}$  is the true value,  $\delta\mathcal{I}_{rc}$  is zero-mean Gaussian noise, and  $\mathbf{R}$  is based on the sensor datasheet<sup>4</sup>.

At each timestep,  $k$ , keypoint detection returns a list of image locations,  $\mathbf{y}_{k,i} = [u \ v]^T$ , with associated covariances,  $\mathbf{Y}_{k,i}$ , where  $u$ , and  $v$  are generally not integers. We use bilinear interpolation of  $\mathcal{I}$  to produce an azimuth/elevation/range measurement,  $\mathbf{z}_{k,i}$ . The uncertainty,  $\mathbf{Q}_{k,i}$ , associated with  $\mathbf{z}_{k,i}$ , is produced by propagation of  $\mathbf{R}$  and  $\mathbf{Y}_{k,i}$  through the interpolation equations.

The standard error term used in VO systems is *reprojection error*—the difference between the observed keypoint location and the predicted keypoint location given the current state estimate. We use a similar error term based on the spherical camera model used to form the intensity images. Each measurement,  $\mathbf{z}_{k,j}$ , corresponds to an observation of landmark  $j$  at time  $k$ . The error term,  $\mathbf{e}_{k,j}$ , is simply

$$\mathbf{e}_{k,j} := \mathbf{z}_{k,j} - \mathbf{h}(\mathbf{x}_{0,k}, \mathbf{p}_0^{j0}), \quad (4)$$

<sup>4</sup><http://www.optech.ca/i3dtechoverview-Illris.htm>



where  $\mathbf{x}_{0,k}$  is a column of state variables for the camera pose at time  $k$  and expressed in the base frame,  $\mathbf{p}_0^{j0}$  is a column of state variables for landmark  $j$  expressed in the base frame, and  $\mathbf{h}(\cdot)$  is our sensor model.

### B. Experimental Data

The data for this experiment was collected at the University of Toronto Institute for Aerospace studies. The Ilris lidar, a Point Grey Research Bumblebee XB3 stereo camera, and a Thales DG-16 Differential GPS unit were mounted on our large field robot as shown in Figure 1. The lidar scans, stereo images, and DGPS points were collected every 0.5 meters over a traverse of approximately 200 meters. As previously mentioned, the Ilris is not a high framerate 3D lidar, so we were limited to this stop-and-scan methodology.

### C. Results

Figure 6 shows a two-dimensional view of our visual odometry localization results with DGPS for groundtruth. The dataset took approximately 13 hours to collect, allowing for a wide range of lighting conditions.

For this dataset, stereo VO outperformed our lidar VO, achieving a total RMS path error of 1.6% of the distance traveled, while the lidar VO achieved a total RMS path error of 2.9% of distance traveled. The lidar VO estimate experienced drift along the  $z$ -axis due to an accumulation of error in the platform's orientation estimate. We believe the stereo orientation estimate was more accurate due to longer feature track lengths<sup>5</sup>. The shorter track lengths may be due to the lidar's spatial undersampling (as discussed below) but an investigation of how to accurately track features through intensity images should be a topic of future work.

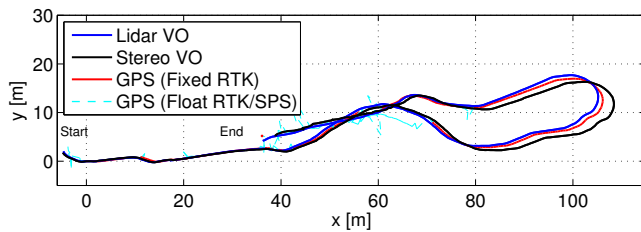


Fig. 6. Two-dimensional view of localization results with DGPS for groundtruth. Due to external interference, our DGPS switched in and out of its fixed Real Time Kinematic (RTK) mode and into its less accurate float RTK and Standard Positioning Service (SPS) mode, causing noticeable jumps in the data. These have been illustrated with a different colour.

Figure 7 shows the number of inlying keypoint matches for both the lidar and the stereo camera over the entire traverse. For most of the traverse, both lidar and stereo achieved a similar number of matches. However, after sunset, we see a significant decrease in matches for stereo, while the number of matches found with lidar remains unchanged. Thus, on average, we were actually able to detect a greater number of inlying keypoint matches using lidar. In addition, as proof

<sup>5</sup>A video showing the feature tracks for both the lidar and camera intensity images can be found at <http://asrl.utias.utoronto.ca/~cmm/videos/>.

of our claim that 125 keypoint matches would be sufficient for motion estimation, we note that the average number of keypoint matches for lidar VO was approximately 121.

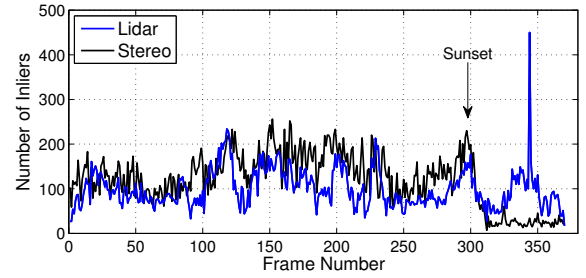


Fig. 7. Number of inlying keypoints versus frame number.

## V. DISCUSSION

Although we have demonstrated that lidar sensors can be used for appearance-based vision methods, there are a number of unique challenges and limitations that need to be addressed. The most obvious and prominent limitation with a lidar sensor is aliasing, which results from spatial undersampling of the scene. Figure 8 shows a zoomed-in view of one of the lidar intensity images taken from our 24 hour experiment, where we see spatial aliasing in the form of a Moire pattern [16]. As a preliminary treatment of aliasing, we employed a simple Gaussian low-pass filter; however, this was not sufficient to reduce the visual distortion. Clearly more appropriate anti-aliasing algorithms need to be investigated, which will be a focus of future work.

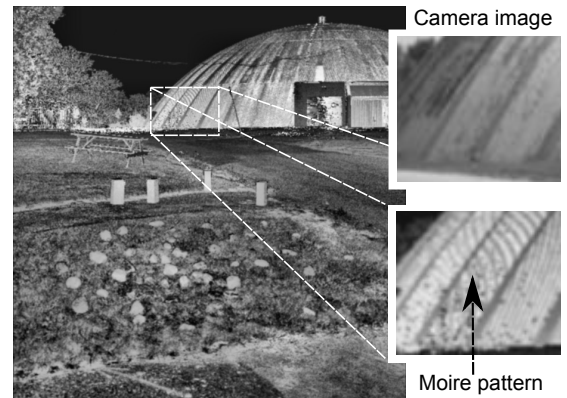


Fig. 8. Zoomed-in view of the dome in one of the lidar intensity images taken from the light sensitivity experiment. The arching light and dark bands on the surface of the dome create what known as a Moire pattern. A camera image of the same segment of the dome is shown for comparison.

Another challenge with active sensors, such as a lidar, is that in non-convex environments, small changes in sensor orientation can result in large range deviations since we only measure point samples in the scene (e.g., objects that have a high angle of incidence with respect to the laser beam or thin objects such as a sign post, can display large range deviations). Figure 9 shows a pixel-wise standard-deviation range image, which was computed by taking all of the range

images we gathered from our light sensitivity experiment and computing the standard deviation of the range values. As can be seen, objects such as the power line, the top of the dome, and the chain link fence represent regions where the range can deviate significantly. Since we use bilinear interpolation to compute keypoint measurements, if one of the neighbouring range values used in our interpolation deviates significantly from the others, it results in a biased range reading. For this reason, we rejected any keypoint measurements where the difference between the maximum and minimum range value was beyond a certain threshold, which occurs mostly at structure boundaries.

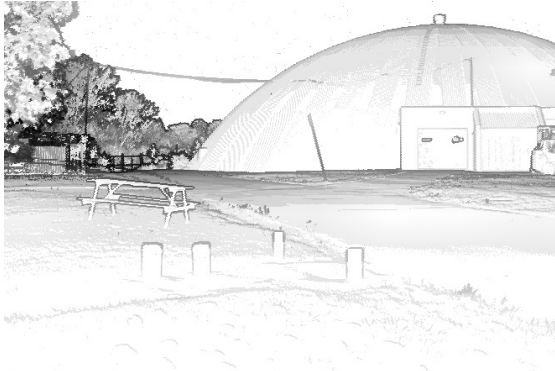


Fig. 9. Standard deviation range image, where black represents the largest deviation. For visual clarity, we applied adaptive histogram equalization to the raw standard deviation image. Note that objects in the distance, such as the top of the dome, the trees and the fence, exhibit the largest range deviations, which are caused by small perturbations in the orientation of the lidar. In other words, for thin objects or surfaces in the distance, it is either a ‘hit’ or ‘miss’ situation. The maximum range deviation was 208m.

A great deal of work still remains to fully realize the potential of active laser sensors. Currently, we are in the process of acquiring a high framerate lidar in order to extend this research to the development of a lidar-based teach and repeat system, similar to our previous work [3]. We anticipate that it will be necessary to provide motion compensation of the lidar images [17]; this is because while capturing lidar images during movement, the laser scan pattern is not visiting every pixel simultaneously.

## VI. CONCLUSION

This paper has demonstrated how active laser-based sensors, such as a lidar, can be used for appearance-based motion estimation techniques, which are predominately accomplished with cameras. Like a camera, a lidar sensor can produce intensity images, allowing the use of sparse keypoint detectors. In our first experiment, which involved taking lidar scans and camera images every half hour for 24 hours, we demonstrated that lidar sensors do in fact exhibit a sensitivity to changes in ambient light. However, even in the worst case scenario of trying to match keypoints from a daytime scan to a nighttime scan, we were still able to match 25% of the keypoints, which is sufficient for motion estimation. In our second experiment, we drove our field robot outdoors and took scans in short increments to simulate a high framerate sensor. Our results showed that the lidar intensity images

are stable enough to implement visual odometry and that the localization estimates are very good when compared with DGPS. This work marks an important step towards enabling appearance-based vision techniques in environments that are not conducive to passive cameras.

## VII. ACKNOWLEDGMENTS

The authors would like to acknowledge the Natural Sciences and Engineering Research Council of Canada and the Canada Foundation for Innovation for providing us with the financial support necessary to conduct our experiments. We would also like to acknowledge Andrew Lambert from the Autonomous Space Robotics Lab (ASRL) for his help in installing the GPS payload and Chi Hay Tong from ASRL for his work on the GPU SURF algorithm.

## REFERENCES

- [1] D. Lowe, “Distinctive image features from scale-invariant key points,” *International Journal of Computer Vision*, vol. 60(2), pp. 91–110, 2004.
- [2] L. Matthies and S. Shafer, “Error modeling in stereo navigation,” *IEEE Journal of Robotics and Automation*, vol. 3, no. 3, 1987.
- [3] P. Furgale and T. Barfoot, “Visual teach and repeat for long-range rover autonomy,” *Journal of Field Robotics, special issue on “Visual mapping and navigation outdoors”*, vol. 27, no. 5, pp. 534–560, 2010.
- [4] K. Konolige, J. Bowman, J. Chen, P. Mihelich, M. Calonder, V. Lepetit, and P. Fua, “View-based maps,” *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 941–957, 2010.
- [5] U. Kretschmer, T. Aaby, M. Thies, and C. Frohlich, “Traffic construction analysis by use of terrestrial laser scanning,” in *Proceedings of the ISPRS working group VIII/2 ‘Laser-Scanners for Forest and Landscape Assessment’*, vol. XXXVI, no. 8/W2, Freiburg, Germany, Oct. 3–6 2004, pp. 232–236.
- [6] C. Dold and C. Brenner, “Registration of terrestrial laser scanning data using planar patches and image data,” in *Proceedings of the ISPRS Commission V Symposium ‘Image Engineering and Vision Metrology’*, vol. XXXVI, no. 5, Dresden, Germany, Sept. 25–27 2006.
- [7] J. Bohm and S. Becker, “Automatic marker-free registration of terrestrial laser scans using reflectance features,” in *Proceedings of the 8th Conference on Optical 3D Measurement Techniques*, Zurich, Switzerland, July 9–12 2007, pp. 338–344.
- [8] M. Fischler and R. Bolles, “Random sample and consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Comm. of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [9] J. Neira, J. Tardos, J. Horn, and G. Schmidt, “Fusing range and intensity images for mobile robot localization,” *IEEE Transactions on Robotics and Automation*, vol. 15, no. 1, 1999.
- [10] J. Guivant, E. Nebot, and S. Baiker, “Autonomous navigation and map building using laser range sensors in outdoor applications,” *Journal of Robotic Systems*, vol. 17, no. 10, pp. 565–583, 2000.
- [11] S. May, S. Fuchs, E. Malis, A. Nuchter, and J. Hertzberg, “Three-dimensional mapping with time-of-flight cameras,” *Journal of Field Robotics*, vol. 26, no. 11–12, pp. 934–965, 2009.
- [12] D. Borrmann, J. Elseberg, K. Lingemann, A. Nuchter, and J. Hertzberg, “Globally consistent 3d mapping with scan matching,” *Journal of Robotics and Autonomous Systems*, vol. 65, no. 2, 2008.
- [13] H. Bay, A. Ess, T. Tuytelaars, and L. Gool, “Surf: Speeded up robust features,” *Computer Vision and Image Understanding (CVIU)*, vol. 110, no. 3, pp. 346–359, 2008.
- [14] G. Sibley, C. Mei, I. Reid, and P. Newman, “Vast-scal outdoor navigation using adaptive relative bundle adjustment,” *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 958–980, 2010.
- [15] K. Konolige, M. Agrawal, and J. Solà, “Large scale visual odometry for rough terrain,” in *Proceedings of the International Symposium on Robotics Research (ISRR)*, 2007.
- [16] G. Oster and Y. Nishijima, “Moire patterns,” *Scientific American*, vol. 208, pp. 54–63, 1963.
- [17] M. Bosse and R. Zlot, “Continuous 3d scan-matching with a spinning 2d laser,” *IEEE International Conference on Robotics and Automation*, pp. 4312–4319, 2009.