

Michael T. Benson

Department of Mechanical Engineering,
Villanova University,
Villanova, PA 19085
e-mail: michael.benson@villanova.edu

Harish Sathishchandra

Research Fellow
Department of Electrical and Computer
Engineering,
Boston University,
Boston, MA 02215
e-mail: harishns@bu.edu

Garrett M. Clayton¹

Mem. ASME
Associate Professor
Director
Center for Nonlinear Dynamics and Control,
Department of Mechanical Engineering,
Villanova University,
Villanova, PA 19085
e-mail: garrett.clayton@villanova.edu

Sean B. Andersson

Mem. ASME
Associate Professor
Department of Mechanical Engineering,
Division of Systems Engineering,
Boston University,
Boston, MA 02215
e-mail: anderss@bu.edu

Compressive Sensing-Based Reconstruction of Lissajous-Like Nodding Lidar Data

In this article, a compressive sensing-based reconstruction algorithm is applied to data acquired from a nodding multibeam Lidar system following a Lissajous-like trajectory. Multibeam Lidar systems provide 3D depth information of the environment, but the vertical resolution of these devices may be insufficient in many applications. To mitigate this issue, the Lidar can be nodded to obtain higher vertical resolution at the cost of increased scan time. Using Lissajous-like nodding trajectories allows for the trade-off between scan time and horizontal and vertical resolutions through the choice of scan parameters. These patterns also naturally subsample the imaged area. In this article, a compressive sensing-based reconstruction algorithm is applied to the data collected during a relatively fast and therefore low-resolution Lissajous-like scan. Experiments and simulations show the feasibility of this method and compare the reconstructions to those made using simple nearest-neighbor interpolation. [DOI: 10.1115/1.4046574]

Keywords: mechatronics, sensors and sensor networks, Lidar

Introduction

For autonomous robots, having range information about the environment and the objects within it is extremely useful, if not necessary, for the robot to navigate and react to its surroundings. The data can be used for localization and/or mapping in the forms of visual odometry [1], SLAM [2], and more [3]. In recent years, Lidar has perhaps become the most widespread way to collect this information in large part due to increasing availability and falling prices as a result of their use in self-driving automobiles. While Lidar configurations come in many different forms, recent systems tend to favor multibeam arrangements, which rotate multiple beams to provide a 3D view of the scene. However, due to limitations on how closely the beams can be spaced, the vertical resolution can be still insufficient for specific applications, as the insufficient vertical resolution can cause objects to “disappear” as they pass between visibility of one beam to another. Even if seen, it is unlikely that detection by a single beam would provide sufficient information to successfully classify the object. Newer, higher resolution Lidar models partially mitigate, but do not completely overcome, this issue.

Before multibeam Lidar systems became more widely available, single-beam planar Lidar units were (and still are, in some applications) used to collect 3D depth information by placing them on a

nodding or spinning platform, allowing the plane in which the Lidar scans to be rotated [4]. Typically, these configurations are driven through a raster scan pattern in which the assembly is rotated in a stair-step pattern between scans of the Lidar. This method results in high resolution in the dimension added due to rotation, but at the expense of significantly increased scan times. In the prior work, two of the authors explored the applicability of raster scan patterns for multibeam Lidar to help solve the “disappearing” problem and how other scan trajectories, such as Lissajous-like scan patterns, can be used to provide a trade-off between scan period and resolution [5]. While Lissajous-like scanning could replicate depth images from the raster scan to some degree, the resulting decrease in the data gathered resulted in smaller objects being more difficult to identify.

While it is often useful to trade-off some resolution for higher frequency sampling, it is typically still desirable to have the highest resolution scan possible to support detection, mapping and SLAM. Inspired by recent results of one of the authors on increasing the imaging rate in atomic force microscopy through subsampling, combined with image reconstruction algorithms to recover high resolution detail [6], in this paper, we explore the use of reconstruction algorithms combined with the Lissajous-like sampling pattern in an attempt to get both high speed and high resolution. The essential idea is to view a high-resolution raster scan image as a “ground truth” image that has been subsampled through the Lissajous-like scan. The final image is formed from the data using a recent optimization algorithm for sparse reconstruction of 3D depth images [7]. This approach differs from common Lidar

¹Corresponding author.

Manuscript received October 22, 2019; final manuscript received February 12, 2020; published online March 4, 2020. Assoc. Editor: Qian Wang.

reconstruction methods, which infer some information about the detected objects from segmentation or classification to determine how to perform reconstruction on the Lidar data, such as in Ref. [8]. Instead, in this work, reconstruction is performed directly on Lidar depth images using compressive sensing (CS) techniques, without performing any other processing on the data.

In this article, experimental results are presented to establish the feasibility of this method. Specifically, Lissajous-like scans that sample less than 10% of the full raster scan data can reproduce the full raster data set with high fidelity when the CS algorithm is applied. To further the advantages of this method at very low sampling levels, the experimental data are subsampled further and then used to reconstruct the full raster data with good results. The reconstructed images are compared both visually and numerically by computing the mean squared error (MSE) between the reconstructions and the ground truth image.

Lissajous-Like Patterns in Nodding Lidar

Much of the work relevant to this section was presented in Ref. [5] and as a result will be presented in minimal depth here. The nodding Lidar system used for this effort is shown in Fig. 1(a). The apparatus features a direct drive system for nodding the Lidar platform and is designed so the distance separation between the Lidar's optical center and the axis of rotation is small. The Lidar used is the Velodyne VLP-16 Puck, which rotates its beams at a rate of 20 Hz, with a horizontal resolution of 900 points per scan over a full 360 deg range and a vertical resolution of 16 equally spaced beams over a 30 deg field of view, which in total produces a maximum output of 14,400 points per scan in single return modes.

The relevant axes and geometry of the assembly is shown in Fig. 1(b). Note that for this work, the angular velocity $\dot{\theta}_l$ that dictates the angle of the beams inside the Lidar θ_l and the incident angle of the beams θ_i are considered fixed; only trajectory design of the nodding motion θ_n about the y-axis is discussed. In this work, the spinning of the beam assembly θ_l rotates according to:

$$\theta_l(t) = -\pi + 2\pi(\dot{\theta}_l t \bmod 1) \quad (1)$$

where $\dot{\theta}_l = 20$ Hz is the fixed rotational velocity of the Lidar, and the incident angles of the 16 beams θ_i are fixed:

$$\theta_i \in \{-15 \text{ deg}, -13 \text{ deg}, \dots, -1 \text{ deg}, 1 \text{ deg}, \dots, 15 \text{ deg}\} \quad (2)$$

Because the parameters of the Lidar are fixed, the distribution of the scan readings is determined by the choice of trajectory for the nodding motion. For a raster scan, the trajectory is a stair-step pattern, with the Lidar nodding in small angle increments between scans of the Lidar, with a trajectory spanning from some arbitrary angle $-A$ to A . For Lissajous-like scanning, the Lidar is nodded with a triangular trajectory between $-A$ and A with frequency ω_n (we call this "Lissajous-like" because Lissajous trajectories are typically sinusoidal in both axes). Both the raster and Lissajous-like trajectories are used to create depth images in this work, with the data from a high-resolution raster scan treated as the "ground truth" and data from low-resolution Lissajous-like scans treated as downsamples of the ground truth data.

Reconstruction of Lidar Images From Subsampled Data

The goal of the reconstruction is to fill in the missing pixels in the depth image (namely, those that were not sampled) with values that depend on the measured data. Perhaps the simplest way to do this is nearest-neighbor (NN) interpolation. Such an approach can be effective if there are sufficient data, but it does not take into account any structure in the image and may therefore produce less than ideal results. An alternative approach is to leverage results from CS. CS is a joint measurement and signal-processing technique, which can produce good (or even exact) reconstructions of signals from significantly fewer measurements than the Nyquist-Shannon sampling theorem requires [9]. At the heart of CS is the assumption of compressibility (or true sparsity) of the signal of interest, that is, when described in an appropriate basis, most of the coefficients are negligible (compressibility) or exactly zero (sparsity).

While there are many reconstruction algorithms, the essential ideas can be understood from the basis pursuit algorithm in which a signal is reconstructed from measurements by solving the ℓ_1 minimization problem:

$$\min \|\eta\|_1 \quad \text{subject to} \quad y = \Phi x = \Phi \Psi \eta \triangleq A \eta \quad (3)$$

where $x \in \mathbb{R}^n$ is the true signal, $y \in \mathbb{R}^m$ is the vector of measurements, Φ is an $m \times n$ measurement matrix, Ψ is an $n \times n$ orthonormal basis for $\mathbb{R}^n$ (often referred to as the *sparsity basis*), and η is the sparse representation of the true signal x in the domain of Ψ . We define the product $\Phi \Psi$ as the $m \times n$ matrix A .

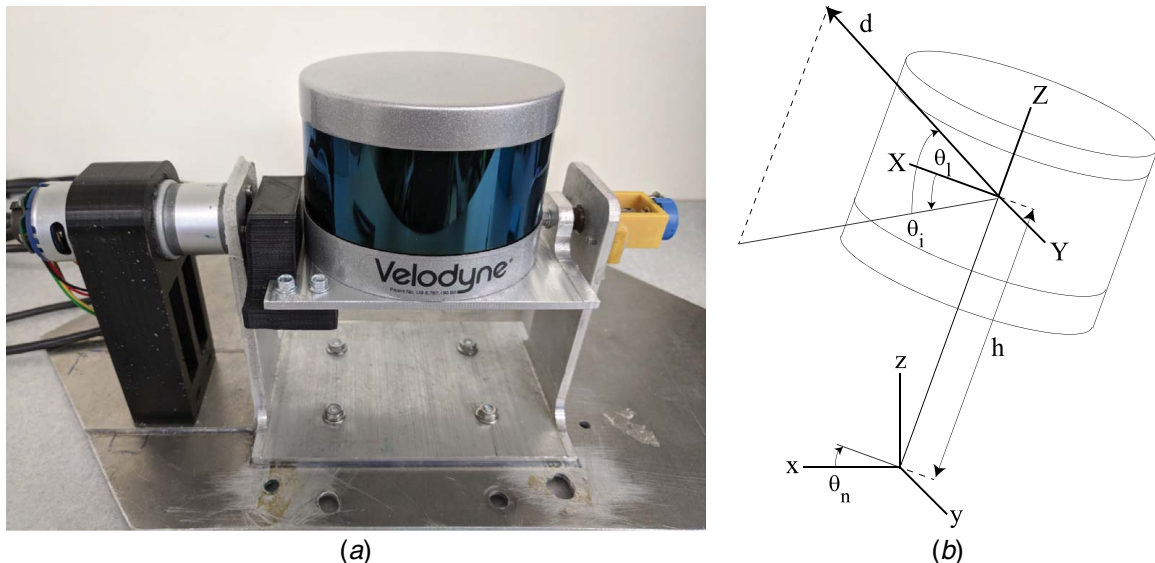


Fig. 1 (a) The Lidar/cradle apparatus used in this research and (b) the relevant axes and notation

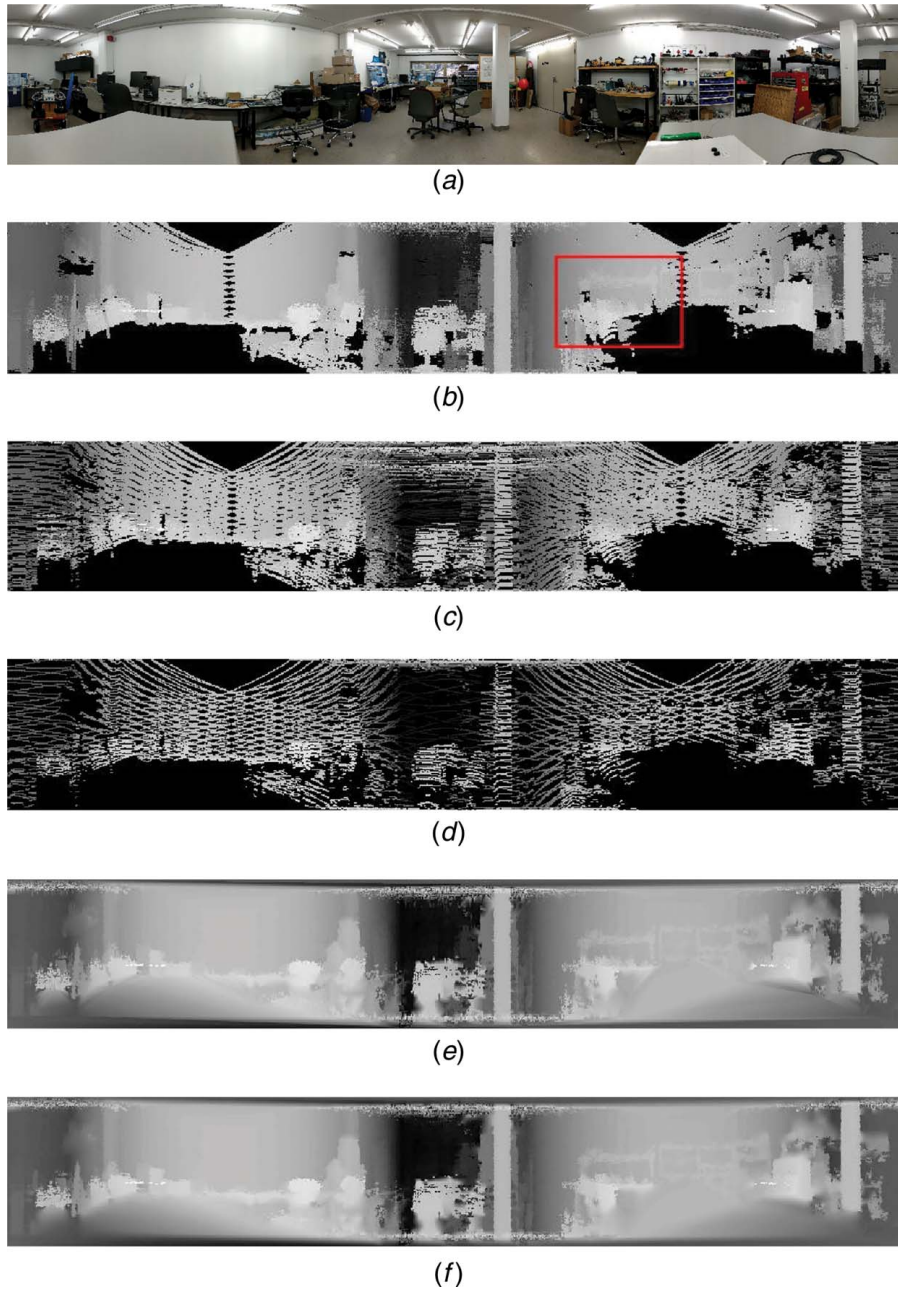


Fig. 2 (a) A color image of the indoor space; (b)–(f) depth images of the same space from different scans and reconstructions: (b) a raster scan with a period of 10 s; Lissajous-like scans with periods of (c) 1/10th (10%) and (d) 1/40th (2.5%) that of the raster scan; and reconstructions based on data from the 2.5% Lissajous-like scan via (e) nearest-neighbor interpolation and (f) compressive sensing-based reconstruction

In the context of reconstruction of a depth image from the nodding Lidar scan, measurements are restricted to values of the pixels of the image that lie on the scanned path. As a result, each row in Φ consists of a single one (indicating which pixel is measured) with all other entries being zero. Thus, the measurement matrix Φ is given by extracting rows from an $n \times n$ identity matrix. This has implications for a property known as the *mutual coherence* [10], generally implying that reconstructions from the data sampled by Φ from solving Eq. (3) will not be exact and in general may have a large error. The metric of mutual coherence is in general improved by using random sampling, and experience in many applications has shown good performance (see, e.g., Refs. [11,12]).

As it turns out, in structured environments such as one finds in indoor environments, the depth profile is often well described as consisting of many planar regions connected by a few “edges.” The sparsity of the data, then, is not with respect to a particular sparsity basis, but rather with respect to these edges. This case was first considered in Ref. [7] where a reconstruction algorithm was created using a cosparsity model. We give a brief description here.

To reflect the fact that we are after a depth profile, let the underlying signal now be denoted as $Z \in \mathbb{R}^{n \times n}$, viewed as a matrix. Since Z is assumed to be formed from (not too many) planar regions, the goal is to find a profile that fits the data while minimizing the number of depth changes in the horizontal, vertical, and diagonal

	2.5%	0.25%	0.25% (Random)
NN			
CS			

Fig. 3 Reconstructions of the highlighted region in Fig. 2(b) from the left column: 2.5% Lissajous-like scan, middle column: 0.25% regularly subsampled data, and the right column: 0.25% randomly subsampled data

directions. These changes can be measured using three different convolutional filters with kernels:

$$D_{xx} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad D_{yy} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad D_{xy} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad (4)$$

Reconstruction of the full depth profile from a collection of measurements y obtained by a measurement matrix Φ comes from solving the optimization problem:

$$\min_Z \|\text{vec}(Z \otimes D_{xx}) + \text{vec}(Z \otimes D_{yy}) + \text{vec}(Z \otimes D_{xy})\|_1 \quad (5)$$

subject to $\|\Phi \text{vec}(Z) - y\|_\infty \leq \epsilon$

where $\text{vec}(\cdot)$ is the vectorization operator, which transforms a matrix into a vector by stacking the columns, $\otimes$ is the convolution operator, and ϵ is a user-defined parameter that accounts for noise in the measurements.

Experimental Imaging and Reconstruction

In this work, an indoor and an outdoor space were imaged by the Lidar. The Lidar followed a nodding trajectory of ± 20 deg, which provided a sufficient field of view to visualize the relevant objects in the space (but not things considered unnecessary, such as the ceiling of the indoor space). The spaces are first scanned by the Lidar undergoing a raster scan trajectory with a period of 10 s. Then, the spaces are scanned by the Lidar with Lissajous-like trajectories with periods of 1 s and 0.25 s. These Lissajous-like scans can be treated as “subsamped” versions of the raster scan, which is treated as the ground truth, with downsampling percentages of 10% and 2.5%, respectively.

Figure 2 shows the indoor space imaged by a color camera and the nodding Lidar. A panoramic color image is presented in Fig. 2(a) to familiarize the reader with the space. The raster scan is shown in Fig. 2(b), and the 1-s and 0.25-s Lissajous-like scans are shown in Figs. 2(c) and 2(d), respectively. The data from the Lidar scans are displayed as depth images, with closer objects being whiter and regions where there are no data from the scan showing as black. Note that, because of how the grid is generated for the depth images, the downsampling percentages of the Lissajous-like scans do not correspond to the percentage of populated pixels in the depth images. This is because of cases where multiple points are associated with the same pixel, commonly along the sides of the scan where measurements are more clustered.

Reconstructions of the indoor Lidar data from the 2.5% Lissajous-like scan in Fig. 2(d) are shown in Fig. 2(e) for the NN interpolation and Fig. 2(f) for the CS-based reconstruction. Both the NN- and CS-based reconstructions are visually quite good. While there are only minor differences between them, in general, edges are (slightly) crisper in the CS-based reconstruction, especially on linear edges like the pillars at the middle and far right of the images.

To explore performance at extremely low sampling, the 2.5% Lissajous-like scan data were further subsampled by discarding some of the data. This was done in two ways: first by keeping only every 10th sample (thus regularly subsampling) and second by randomly selecting 10% of the sample points, both resulting in a total subsampling of 0.25% from the raster scan. (In practice, of course, one would not discard data but

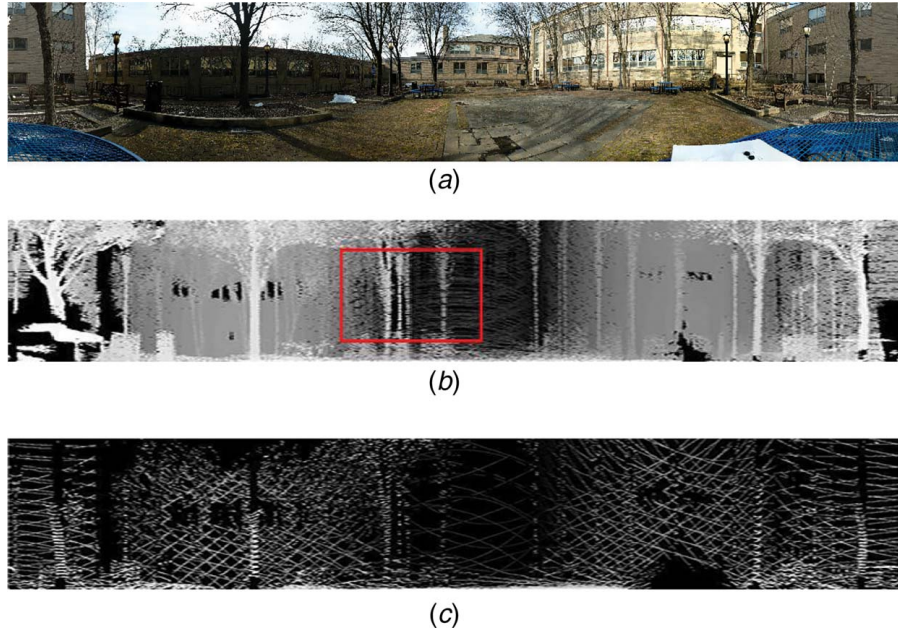


Fig. 4 (a) A color image of the outdoor space; depth images of the same space from (b) a raster scan with a period of 10 s and (c) a Lissajous-like scan with a period of 1/40th (2.5%) that of the raster scan

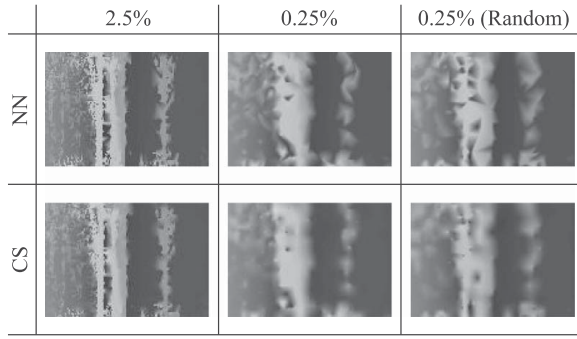


Fig. 5 Reconstructions of the highlighted region in Fig. 4(b) from the left column: 2.5% Lissajous-like scan, middle column: 0.25% regularly subsampled data, and the right column: 0.25% randomly subsampled data

rather choose to only acquire the limited set.) The NN- and CS-based reconstructions were performed again on these 0.25% downsamples.

To compare all the reconstructions, we consider the focus area bound by a red rectangle in Fig. 2(b) containing a workbench and a chair. Each reconstruction of this region from the 2.5% Lissajous-like scan, as well as from the 10% downsamples of the 2.5% Lissajous-like scan, is shown in Fig. 3. Qualitatively speaking, the appearance of the bench and the chair seems to have more detail, but are more prone to noise around the edges, in the 2.5% reconstructions in Fig. 3 (left column). In the middle and right columns of Fig. 3, both the nearest neighbor and CS reconstructions of the 0.25% are remarkably good given the extremely low sampling, but clearly are degraded relative to the reconstructions from the 2.5% data. For the 0.25% reconstructions, the same relationship between the CS and NN reconstruction is observed as earlier: the edges tend to be slightly more crisp in the CS reconstruction. In addition, for the CS reconstructions, the randomly subsampled image is visually superior to the regularly subsampled version despite the same level of subsampling. Interestingly, the nearest-neighbor interpolation seems to have darkened the randomly subsampled 0.25% scan during the reconstruction process, while the compressive sensing-based method retained the original intensity (and hence, depth values).

Figure 4 shows the outdoor space imaged by a color camera and depth images from the raster and 2.5% Lissajous-like scans. While the full reconstruction results are not shown here, the reconstructions of the 2.5% subsampled data behave similarly as the indoor case, in that both algorithms have comparable performance but the CS reconstruction tends to have sharper edges. The reconstructions for the region highlighted in Fig. 4(b) are shown in Fig. 5. This is an especially difficult region for the reconstruction techniques to resolve as the trees are very close to each other but not touching. As before, the NN and CS techniques applied to the 2.5% Lissajous-like scan data, shown in Fig. 5 (left column), have comparable performance although the edges of the trees are slightly more distinct in the CS reconstruction. Further, downsampling in Fig. 5 (middle and right columns)

significantly degrades the reconstruction performance, with the trees being merged into one object in both reconstructions of the regularly subsampled data. When reconstructing the randomly subsampled data, the CS method almost manages to resolve the two middle trees as separate objects, while the NN interpolation resolves them into a single shape with a few holes.

Discussion

In order to compare the reconstruction results numerically, the MSE between the reconstructed image I_r and the ground truth (raster) image I is computed,

$$\text{MSE} = \frac{1}{n} \sum_{p \in \mathbf{p}} (\text{vec}(I_r)[p] - \text{vec}(I)[p])^2 \quad (6)$$

where $\mathbf{p} = \{p_1, \dots, p_n : \text{vec}(I)[p_i] > 0 \forall p_i\}$ is the set of indices for which the corresponding pixels of I are nonzero. The MSE is then divided by the mean of the nonzero points in the ground truth image to scale the results to a reasonable range:

$$\text{MSE}' = \frac{n \cdot \text{MSE}}{\sum_{p \in \mathbf{p}} \text{vec}(I)[p]} \quad (7)$$

The values of MSE' are presented in Table 1 for reconstructions from the 2.5% and 0.25% data. The MSE' results are shown not only for the full images (indoor full, outdoor full) but also for the regions of the images highlighted in Figs. 2(b) and 4(a). Both reconstructions from the 2.5% (nonsampled) scan have comparable performance, except the outdoor trees subsection, in which the CS technique resolved the edges of the trees significantly better than NN for that particular region. The results for the reconstructions start to diverge when the scans are subsampled further. When the 2.5% data are regularly subsampled, the nearest-neighbor interpolation outperforms the compressive sensing-based reconstruction by a small margin (5–10%) indoors and by a smaller margin outdoors. However, when the 2.5% data are randomly subsampled instead, the compressive sensing-based reconstruction outperforms the nearest-neighbor interpolation by a large margin (25–35%) indoors, and a smaller performance improvement ($\approx 10\%$) outdoors for the trees region outdoors, but not for the full outdoor image.

In general, it appears that the CS reconstruction has comparable or better performance than the NN technique for indoor environments, with significant improvement when the data are randomly sampled. The CS reconstruction also typically appears to have better performance in regions with straight edges, such as the pillars in the indoor space or the highlighted region of the outdoor space. This is likely a result of the choice of convolutional kernels, which were designed to fit the data profile while minimizing the number of depth changes in the horizontal, vertical, and diagonal directions. This indicates that perhaps appropriate kernels could be designed for other environments with known, but structure that is more complicated than simple planes and straight lines.

In all cases but the full outdoor image, the CS technique performed significantly better on randomly subsampled data than regularly subsampled data, highlighting that randomness is important in subsampling. In fact, it has been shown that random sampling matrices are maximally incoherent with any fixed sensing matrix [13] and should thus be expected to outperform deterministic sampling in the general setting.

Conclusions

The use of CS reconstruction algorithms for Lissajous-like nodding Lidar was presented. This method allows for very small amounts of data and thus very small data collection times to

Table 1 MSE' values of depth image reconstructions

Subsampling reconstruction	2.5%		0.25%		0.25% (Random)	
	NN	CS	NN	CS	NN	CS
Indoor full	27.43	27.29	16.15	17.01	20.49	15.14
Indoor bench	13.96	13.9	13.22	14.29	18.55	12.54
Outdoor full	15.77	15.31	15.99	15.65	15.92	15.9
Outdoor trees	23.11	20.78	23.1	22.49	20.84	18.97

produce high fidelity depth data. Experimental and simulation results show the feasibility of this approach in both indoor and outdoor scenarios.

Acknowledgment

The work of Sathishchandra and Andersson was supported in part by the NSF through CMMI-1562031. The work of Benson and Clayton was supported in part by the NSF through OISE-1658696.

References

- [1] Fraundorfer, F., and Scaramuzza, D., 2012, "Visual Odometry: Part II: Matching, Robustness, Optimization, and Applications," *Rob. Automat. Mag., IEEE*, **19**(2), pp. 78–90.
- [2] Kohlbrecher, S., Von Stryk, O., Meyer, J., and Klingauf, U., 2011, "A Flexible and Scalable Slam System With Full 3d Motion Estimation," 2011 IEEE International Symposium on Safety, Security, and Rescue Robotics, Kyoto, Japan, Nov. 1, IEEE, pp. 155–160.
- [3] Geller, D., 2007, "Orbital Rendezvous: When Is Autonomy Required?," *J. Guid. Control Dyn.*, **30**(4), pp. 974–981.
- [4] Desai, A., and Huber, D., 2009, "Objective Evaluation of Scanning Ladar Configurations for Mobile Robots," IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, Oct. 10, IEEE, pp. 2182–2189.
- [5] Benson, M., Nikolaidis, J., and Clayton, G. M., 2018, "Lissajous-Like Scan Pattern for a Nodding Multi-Beam Lidar," ASME 2018 Dynamic Systems and Control Conference, Atlanta, GA, Sept. 30, American Society of Mechanical Engineers, p. V002T24A007.
- [6] Luo, Y., and Andersson, S. B., 2019, "A Continuous Sampling Pattern Design Algorithm for Atomic Force Microscopy Images," *Ultramicroscopy*, **196**, pp. 167–179.
- [7] Ma, F., Carlone, L., Ayaz, U., and Karaman, S., 2016, "Sparse Sensing for Resource-Constrained Depth Reconstruction," 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea, Oct. 9, IEEE, pp. 96–103.
- [8] Sampath, A., and Shan, J., 2009, "Segmentation and Reconstruction of Polyhedral Building Roofs From Aerial Lidar Point Clouds," *IEEE Trans. Geosci. Remote Sens.*, **48**(3), pp. 1554–1567.
- [9] Candès, E. J., 2006, "Compressive Sampling," International Congress of Mathematicians, Madrid, Spain, Aug. 22, pp. 1433–1452.
- [10] Candès, E. J., and Romberg, J., 2007, "Sparsity and Incoherence in Compressive Sampling," *Inverse Probl.*, **23**(3), pp. 969–985.
- [11] Braker, R. A., Luo, Y., Pao, L. Y., and Andersson, S. B., 2018, "Hardware Demonstration of Atomic Force Microscopy Imaging Via Compressive Sensing and μ -Path Scans," American Control Conference (ACC), Milwaukee, WI, June 27, pp. 6037–6042.
- [12] Ma, F., Cavalheiro, G. V., and Karaman, S., 2019, "Self-Supervised Sparse-to-Dense: Self-Supervised Depth Completion From LiDAR and Monocular Camera," 2019 International Conference on Robotics and Automation (ICRA), Montreal, Canada, May 20, pp. 3288–3295.
- [13] Baraniuk, R., Davenport, M., DeVore, R., and Wakin, M., 2008, "A Simple Proof of the Restricted Isometry Property for Random Matrices," *Construct. Approximation*, **28**(3), pp. 253–263.