REMOTE-SENSED LIDAR USING RANDOM IMPULSIVE SCANS

Juan Castorena and Charles D. Creusere
Advising Professor: Dr. Charles D. Creusere
Department of Electrical and Computer Engineering
New Mexico State University
Las Cruces, NM, 88003, USA

Abstract

Third generation full-waveform (FW) LIDAR systems image an entire scene by emitting laser pulses in particular directions and measuring the echoes. Each of these echoes provides range measurements about the objects intercepted by the laser pulse along a specified direction. By scanning through a specified region using a series of emitted pulses and observing their echoes, connected 1D profiles of 3D scenes can be readily obtained. This extra information has proven helpful in providing additional insight into the scene structure which can be used to construct effective characterizations and classifications. Unfortunately, massive amounts of data are typically collected which impose storage, processing and transmission limitations. To address these problems, a number of compression approaches have been developed in the literature. These, however, generally require the initial acquisition of large amounts of data only to later discard most of it by exploiting redundancies, thus sampling inefficiently. Based on this, our main goal is to apply efficient and effective LIDAR sampling schemes that achieve acceptable reconstruction quality of the 3D scenes. To achieve this goal, we propose on using compressive sampling by emitting pulses only into random locations within the scene and collecting only the corresponding returned FW signals. Under this framework, the number of emissions would typically be much smaller than what traditional LIDAR systems require. Application of this requires, however, that scenes contain many degrees of freedom. Fortunately, such a requirement is satisfied in most natural and man-made scenes. Here, we propose to use a measure of rank as the measure of degrees of freedom. To recover the connected 1D profiles of the 3D scene, matrix completion is applied to the tensor slices. In this paper, we test our approach by showing that recovery of compressively sampled 1D profiles of actual 3D scenes is possible using only a subset of measurements.

Keywords: LIDAR, full-waveform, compressive sensing, matrix completion, low-rank approximations

1 Introduction

1.1 LIDAR Background

The advent and usage of third generation pulsed LIDAR systems has increased interest in new sampling, processing, transmission and compression algorithms to suit the requirements implied by these systems. The operation principles of these systems consist in general, on projecting an energy pulse into a scene and measure the reflected full-waveform (FW) signal. Each of these reflected signals provides a range measurement of the intercepted objects along a specified direction and the combination of them results in the range map of the region. By scanning through a specified region by series of emitted pulses and observing their reflected FW signals, connected 1D profiles of 3D scenes can be readily obtained. These waveforms can be further processed to obtain 3D point clouds representing the topography of the sampled region. Recently, usage of third generation LIDAR systems has gained popularity over first and second generation LIDAR's [13]. The advantages of these systems over conventional first and second generation LIDAR's storing 3D point clouds instead of contiguous full 1D profile waveforms is that the shape of the waveforms provides additional insight into the surface structure that can be used for improved characterization and classification. This additional information has proven helpful in applications such as forest analysis and modeling and in 3D point cloud classification in urban areas [12].

There are several problems affecting the performance of this tool, however. One of which is that typical LIDAR range maps are represented by large amounts of data, creating significant challenges for the storage, processing and transmission even in the case of 1st and 2nd generation LIDAR's storing point clouds. Such a problem is aggravated even more when storing FW signals as in 3rd gen. LIDAR. In addition, there are noise sources (e.g. airplane movement [15]) affecting the precision of the trajectories followed by the scanning pattern which results in an "unstructured" data cloud. For this reason, storage of the reference positioning where each pulse was transmitted is required increasing memory limitations.

1.2 Compression Approaches

To resolve these issues, a number of LIDAR compression approaches have been developed in the literature [8], [11]. These, however generally require the initial acquisition of large amounts of data only to later discard most of it by exploiting redundancies. More recent approaches such as compressive sensing [3], rely on projecting data into a low dimensional subspace by random sub-gaussian linear transformations. Under this framework, a signal can be sampled at a rate much lower than required by the Nyquist/Shannon theorem under a signal sparsity constraint. In [6], the authors applied compressive sampling to 1st generation LIDAR's returning point clouds which can be represented by 2D matrices. However, practical implementations of these algorithms to connected 1D profiles of the 3D scene are at this point unfeasible. The main limitations are the high computational complexity involved in the reconstruction algorithms, that a sparsity basis is generally unknown and that projecting returning waveforms into random spaces prior to digitization is not feasible with current optical LIDAR components. In [4], the authors compressively sampled connected 1D profiles of 3D scenes successfully by considering FW signals as signals of finite rate of innovation.

In that research, connected 1D profiles of the 3D scene where recovered very precisely by uniformly sampling FW signals at rates much lower than what Nyquist/Shannon dictates. Although, high quality reconstructions where achieved with high compression ratios the approach only explores temporal redundancies.

In this paper, we focus on compressive sensing by exploring the spatial redundancies that may be present in connected 1D profiles of 3D scenes. Here, we explore this venue by considering surfaces with many degrees of freedom. Such an assumption is valid for many natural and man made scenes. The measure of degrees of freedom of a surface is based on a rank measure. The lower the rank the lower the number of measurements required. Based on this feature, we propose to scan randomly and impulsively across different portions of the scene according to a uniform distribution. Each of these collected waveforms consists of a time-resolved waveform describing ranges along a path. Here, we analyze the conditions under which good quality reconstructions can be achieved based on the sub-sampled surface. Such conditions will determine, the density of random impulsive scan samples required to achieve some error bound. To achieve this, we use the theoretical results obtained in [5] and the concepts from [14] to determine the required sub-sampling densities. Reconstruction of the connected 1D profiles of the 3D scene is achieved via matrix completion [2]. In general, we apply the reconstruction algorithms used in matrix completion by slicing LIDAR datasets into a contiguous sequence of matrices each representing a snapshot of the FW signals at certain time instant.

1.3 Outline

In section 2 we briefly review the problem of matrix completion, state the problem mathematically and describe the conditions under which recovery is possible with high probability. Section 3 describes the sensing scheme we propose, the model of the sensing scheme, the conditions that guarantee recovery and the procedure used to reconstruct connected 1D profiles of the 3D scene. The reconstruction procedure consists on applying matrix completion to a tensor. Section 4 presents experimentation showing the reconstruction of a dataset obtained from a real LIDAR system. Finally, section 5 discusses the obtained results, and describes the advantages and limitations of our approach found in this research.

2 MATRIX COMPLETION

In general, the problem of matrix completion consists on recovering a matrix from an incomplete observation of its entries [2]. Say, we want to recover a matrix $A \in \mathbb{R}^{M\times N}$ based on $m = |\Omega| \ll MN$ observation of its entries (the set Ω is a set containing the indices of the observed entries). In other words, we are given an incomplete matrix X given by the entries

$$X_{i,j} = \begin{cases} A_{i,j}, & (i,j) \in \Omega \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

and the problem is to recover A by only observing X. This operation is modeled by application of the projection operator $P_\Omega : \mathbb{R}^{M\times N} \to \mathbb{R}^{M\times N}$ (i.e., $X = P_\Omega(A)$) which projects into the subspace of matrices that vanish outside of Ω. Surprisingly, turns out that a matrix A

can be recovered if the matrix is of low rank and if the number of observed entries satisfies a number dependent upon this number [2]. The solution to this problem is obtained by finding the matrix $\hat{A}$ with minimum rank that satisfies the observations $X_{i,j}$. Unfortunately, this problem is highly nonlinear. However, in [2] it was found that the problem can be recasted as a convex optimization problem involving minimization of the nuclear norm as

$$\begin{array}{ll} \text{minimize} & \|X\|_* \\ \text{subject to} & X_{i,j} = A_{i,j},\ (i,j) \in \Omega \end{array} \tag{2}$$

where $\|X\|_*$ denotes the nuclear norm given by the sum of the singular values as $\|X\|_* = \sum_{k=1}^{n} \sigma_k(X)$. Here, $n = \min(M, N)$ and $\sigma_k(X)$ denotes the kth largest singular value of X. Equation (2) describes the problem when there are no noise sources. We know however, that this situation is not practical in real systems composed of non-ideal components. Based on this, we model the observations of the matrix entries obtained by the system corrupted with noise $Z_{i,j}$ by

$$X_{i,j} = \begin{cases} A_{i,j} + Z_{i,j}, & (i,j) \in \Omega \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

It was found in [1] that when perfect noiseless recovery occurs, then matrix completion is stable under small perturbations. The bound on the estimation errors is given in [1] by

$$\|A - \hat{A}\|_F \geq 4\sqrt{\frac{C_p n}{p}}\delta + 2\delta \tag{4}$$

for some $\delta > 0$ and where $C_p = 2 + p$ and $p = m/(MN)$. The operator denoted by $\|\|_F$ is referred to as the Frobenious norm. This quantity is the matrix equivalent of the l_2 norm. In fact the Frobenious norm of a matrix is defined here as

$$\|A\|_F = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} A_{i,j}^2. \tag{5}$$

where $A_{i,j}$ is the element of the matrix A indexed by i, j. Note in equation (4) that if the constant δ is small given a number of measurements m, then $\hat{A}$ is close to A in the Frobenious norm sense. A small Frobenious norm error can be interpreted as estimated entries that are close to the true complete entries for all (i, j). Therefore, based on the mentioned statements and on the results given in [1], stable reconstruction is possible with the optimization problem given in equation (2) with probability $1 - cn^{-3}$ when

$$m \geq Cn^{5/4} r \log n \tag{6}$$

for positive constants c and C. Here, the number r denotes the rank of the matrix A to be reconstructed. Because this number is not available prior to sampling, then we cannot determine the lowest number of measurements exactly. However, this number can be estimated based on prior knowledge of the matrix to be recovered.

3 LIDAR SCANNING

3.1 Scanning a Scene

The proposed system emits pulses in random locations across the scene to be sampled and collects the corresponding echoes and stores them as FW signals. These locations are randomly selected (without replacement) from a uniform distribution which spans the entire area covered by the scene to be sampled. At each of the chosen locations, a laser pulse of finite duration is transmitted to the surface and the reflected waveform is measured. The set of locations that can be probably sampled depends here on several factors for example the area to be covered and the laser spot or footprint size. Here, we assume no overlap is present between the footprint size of the emitted pulses.

The entire surface to be sampled is rasterized to form uniform sized squares or pixels in a spatial cartesian coordinate system. Each of these subregions or pixels conforms the possible locations that can be sampled. Since we are interested in sub-sampling, then, locations along the rasterized surface contains missing observations or pixels. To illustrate the random emission of pulses across a region see Figure 1. As illustrated in the figure,

Figure 1: Airborne LIDAR randomly emitting pulses across a region.

pulses are emitted into random locations of a surface from a airborne/remote location. A rotating mirror directs pulses along the perpendicular direction of airborne/remote travel and measures its echoes as a FW signal. The entire surface is then scanned as the airplane traverses the entire region. Note that in random scan impulsive sampling, the location of pulse emissions is random. Also note that with the proposed sensing approach missing information is present only spatially but not temporally (i.e., only missing fibers is present and not missing temporal information along the fibers).

3.2 Model

Based on our sampling approach, we model the collection of a complete connected 1D profile of the 3D scene as a tensor denoted by calligraphic $\mathcal{A}$ with coordinates along the first two dimensions of size $n_1 \times n_2$ describing the spatial locations and with the third dimension of size N describing a normalized time-resolved waveform. We assume the first two dimensions are of the same nature (although this event may not occur in practice because the process involved in scanning is distinct in each direction). A slice of the tensor along a dimension say k (i.e. temporal dimension) is a 2D matrix containing elements in the two other dimensions

(i, j). Each slice is denoted here by the matrix A_k at slice index k. A fiber is referred to in here as the FW signal containing the ranging information in the location obtained by fixing the two spatial directions. Each of these fibers is denoted by $\mathcal{A}_{i,j,:} \in \mathbb{R}^{1\times 1\times N}$. We assume, each pulse of the complete tensor $\mathcal{A}$ is emitted uniformly at equidistant locations. In practice this is not true however, and some interpolation is typically applied to the unstructured cloud of points. Fortunately, we fold the interpolation step into the reconstruction by using random impulsive scanning schemes and matrix completion.

Under the perspective given in section 3.1, emitting pulses into random locations and collecting the corresponding echoes is modeled as the problem of missing fibers within the tensor of measurements. Because the LIDAR system introduces noise, then we model the collected noisy measurements as

$$\mathcal{X}_{i,j,:} = \begin{cases} \mathcal{A}_{i,j,:} + \mathcal{Z}_{i,j,:}, & (i,j) \in \Omega \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where the index set $\Omega \subset \{(1,1),(1,2),...,(n_1,n_2)\}$ is selected randomly according to a uniform distribution. The noise $\mathcal{Z}_{i,j,:}$, is assumed here to be Gaussian and independent for all indices i, j. Such an assumption is valid because each pulse emission and corresponding measurement is independent from each other. In addition, this noise is also independent with respect to the time of acquisition indicated by the index k.

3.3 LIDAR Data Reconstruction

The problem becomes that of recovering tensor $\mathcal{A}$ based on the measurements $\mathcal{X}$. For this purpose, we use the procedure of matrix completion described in section 2. To fit the matrix completion framework, we slice the tensor to be recovered at fixed acquisition times indexed by k. In such a way we treat a tensor as a set of matrices A_k and reconstruct each of these individually. All of these matrices will have missing fibers (i.e., FW signals) at fixed indices $(i, j) \notin \Omega$. Viewing the problem in such a manner allows one to reconstruct tensors $\mathcal{A}$ by slicing and using matrix completion independently on each of the slices. In other words, the problem becomes that of recovering $\mathcal{A}$ using the optimization problem

$$\begin{array}{ll} \text{minimize} & \|X_k\|_* \\ \text{subject to} & X_{i,j,k} = A_{i,j,k},\ (i,j) \in \Omega \end{array} \tag{8}$$

independently for all $k \in (1, N)$ as described in section 2. In a similar fashion to equation (2), (8) recasts the problem of minimizing the rank of a matrix X_k subject to the observations into a convex problem invoking the nuclear norm. This problem finds the correct solution with high probability if the matrix X_k is of low rank or if it can be approximated by a low-rank matrix. Since in our tensor problem Ω is fixed to all slices, then we propose that the density or the number of observations required to recover tensor $\mathcal{A}$ depends upon the quantity $r_{max} = \max_k \text{rank}(A_k)$. Choosing such a value in determining the bound on the number of measurements, guarantees the recovery of the tensor high probability. Given this, the number of measurements required is

$$m \geq C n^{5/4} r_{max} \log n \tag{9}$$

for positive constants c and C and where $n = \max(n_1, n_2)$. If this bound is satisfied, then the optimization problem in equation (8) recovers tensor $\mathcal{A}$ with probability at least $1 - cn^{-3}$. Therefore, a number of pulse emissions into random locations with the corresponding echo measurements satisfying the bound in (9) will recover the complete connected 1D profiles of the 3D scene if the degrees of freedom of the underlying scene is high. In other words, if r_{max} corresponding to the maximum rank of the slices A_k is small, then we can compressively sense the 3D scene and recover it with high probability.

Although r_{max} is not available exactly prior to sampling, we can estimate it by characterizing the distinct surface types corresponding to forests, urban areas, etc. In fact, [7] characterize surfaces based on characterization of the waveform returns. The authors, successfully characterized via simulation the rank of surfaces of distinct complexities by characterizing waveforms through their innovation rate. By surface complexity we refer to surfaces with higher or lower randomness.

4 EXPERIMENTATION

4.1 LIDAR Dataset

The LIDAR dataset obtained from Navair China Lake, CA was collected from the VISSTA ELT LADAR system. The shot rate of the system is of 20 Khz (i..e, pulse emission rate). Each time a pulse of 1.5 ns of duration at full width half maximum is transmitted, the echoes are measured at a sampling rate of 2Ghz and quantized using an 8-bit A/D converter. The dataset used in this research was collected by imaging a pickup truck through a chain link fence both positioned perpendicular to the pulse transmission path. To illustrate this more clearly, Figure 2 shows the 3D point cloud resulting from processing the waveforms collected by the system. The total number of pulse emissions and corresponding measurements used in this dataset are of 31,626 representing a total of 126×251 along the vertical and horizontal directions, respectively. Processing of this dataset was all performed in Matlab.

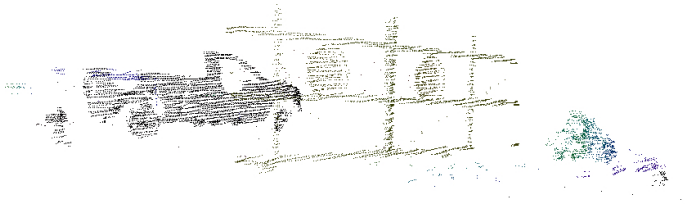


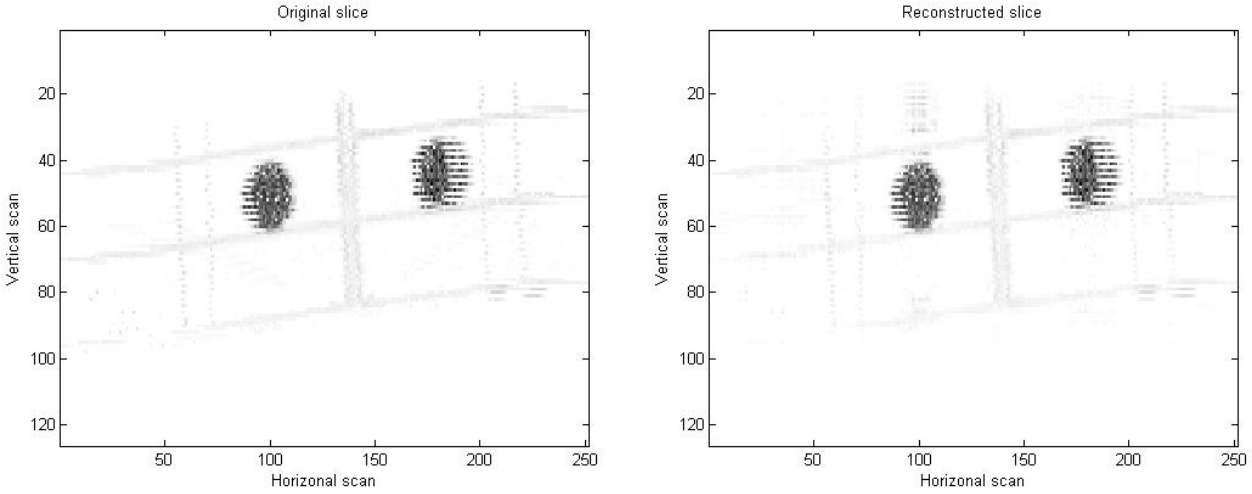
Figure 2: An example of the LIDAR point cloud.

4.2 Scanning and Reconstruction Results

The unstructured connected 1D profiles of the 3D scene was rearranged into a tensor $\mathcal{A}$ of size $126 \times 251 \times 332$ under various assumptions. This tensor represents the complete measurements one would obtain using conventional sampling theorems. To simulate a random impulsive scan as proposed in section 3, we randomly select the subset Ω. The indices in this subset

are drawn according to a uniform distribution from the set $\{(1,1),(1,2),...,(126,251)\}$. The cardinality of the set Ω is given by the number of measurements given in (9) (i.e., $m = |\Omega|$). From prior knowledge of the set, we know that the maximum rank is given as $r_{max} = 116$. However, we assume $r_{max} = 40$ for the purpose of sub-sampling. Using this number and equation (9) with $0 < C < 1$ we chose a total of $m = 8,402$ pulse emissions for sensing the 3D scene. This corresponds to a compression ratio (CR) of $CR = [1 - \frac{m}{n_1 n_2}] \cdot 100\% = 73.44\%$. Based on the chosen set Ω, we construct the incomplete tensor of measurements $\mathcal{X}$ as given in equation (7). The tensor $\mathcal{A}$ is then reconstructed using the OptSpace algorithm described in [10] with an implementation available for Matlab in [9].

An illustration of the matrix completion algorithm applied to our tensor dataset as described in section 3.3 is given for a slice at $k = 134$ in Figure 3a and 3b for the original and reconstructed surfaces, respectively. Here, a white and black color corresponds to zero and one, respectively and the values in between describe the intermediate grayscale levels. Note that the shape of the slice is perceptually well preserved by using only 26.56% of the total number of pulses one would typically emit. To measure the error of the tensor we use



(a) Original slice (b) Reconstructed slice

Figure 3: Reconstruction of a slice of the connected 1D profiles of the 3D scene.

the normalized Frobenious norm (NFN) as a measure of error between the complete and the reconstructed tensors. This measure is given as

$$\text{NFN} = \|\hat{\mathcal{A}} - \mathcal{A}\|_F / (N n_1 n_2). \tag{10}$$

where the Frobenious norm of a tensor is given as $\|\mathcal{A}\|_F = \sum_{k=1}^{N} \|A_k\|_F$. This quantity is given here by the sum of the individual Frobenious norm of its forming slices. Using this measure, the resulting NFN of the error of the truck dataset using our reconstruction approach is 9.9e-3. This small number, exemplifies that reconstruction of a 3D scene is possible using a random location pulse emission LIDAR scheme which compressively senses across the entire desired scenery.

5 DISCUSSION

In this paper, we show that recovery of a 3D scene is possible using a random location pulse emission LIDAR scheme. One advantage of this approach along the already mentioned benefits is that implementation can be achieved with existing third generation LIDAR systems with no extra components required. Unfortunately, in this research we find several limitations of our approach. One is that connected 1D profiles of the 3D scene are in general near sparse containing many zero entries as exemplified in Figure 3. This fact implies that the rank of each of the slices is in general large, thus increasing the number of measurement requirements as described by equation (9). This problem could be alleviated by introducing a low-pass filter blurring each FW-signal received. Another approach, to alleviate this situation could be to use laser pulses of large foot-print size. However, this approach would imply lower resolution reconstructions. Near sparse connected 1D profiles of the 3D scene also imply that random impulsive scanning could contain missing information with no hope of recovery. In general, the higher the sparsity, the larger the constraints on possibly reducing the number of pulse emissions and corresponding FW measurements. In contrast, the lower the sparsity of the slices of the 1D profile of the 3D scenes the higher the capabilities to achieve good quality reconstructions with fewer number of observations.

Another potential drawback of this approach is that in general, the reconstruction algorithms are of high computational complexity requiring long time periods for recovering the LIDAR dataset. In general, taking into account spatial correlations that may be present in such large datasets is highly complex requiring a great deal of operations and memory resources. One possible approach that could be followed here is to subdivide the scene to be scanned into statistically consistent regions as mentioned in [7]. These consistent regions within the scene could be determined with analysis of the complexity of returning FW signal as also proposed in [7]. This approach could reduce the computational and time complexity involved in the reconstructions at the cost of compression efficiency. This efficiency would be reduced by taking advantage only of the correlations within the sub-regions of the scene. Fortunately, correlations between distinct sub-regions would be small in such near sparse scenarios present in connected 1D profiles of the 3D scene.

6 Conclusion

In this paper we deal with the problem of efficient and effective LIDAR sampling of connected 1D profiles (FW signals) of 3D scenes. The proposed compressive sensing scheme is based on emitting laser pulses into random locations spanning the entire region to be imaged. Here we show via theory and experimental work with real LIDAR data that recovery is possible given that the underlying 3D scenes contain many degrees of freedom (similar to sparsity for 1D signals) and that a certain number of measurements is collected. It was found that the number of measurements is lower than what traditional LIDAR systems require given surfaces with a small number of degrees of freedom, thus achieving efficient sampling. The measure of degrees of freedom we proposed here was the maximum rank of all slices conforming the complete tensor of measurements.

References

[1] E.J. Candes and Y. Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2009.

[2] E.J. Candes and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, 2009.

[3] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, March 2008.

[4] J. Castorena and C.D. Creusere. Compressive sampling of lidar: Full-waveforms as signals of finite rate of innovation. In *20th European Signal Processing Conference*, volume Accepted paper, Bucharest, Romania, August 2012.

[5] J. Castorena and C.D. Creusere. Random impulsive scan for lidar sampling. In *IEEE International Conference on Image Processing*, volume Accepted paper, 2012.

[6] J. Castorena, C.D. Creusere, and D. Voelz. Modeling lidar scene sparsity using compressive sensing. In *IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, pages 2186–2189. IEEE, 2010.

[7] J. Castorena, C.D. Creusere, and D. Voelz. Using finite moment rate of innovation for lidar waveform complexity estimation. In *IEEE Conference Record of the Fourty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 608–612. IEEE, 2010.

[8] J. Dagher, M. Marcellin, and M. Neifeld. Efficient storage and transmission of ladar imagery. *Applied Optics*, 42(35):7023–7035, December 2003.

[9] R.H. Keshavan. http://www.stanford.edu/ raghuram/optspace/code.html.

[10] R.H. Keshavan, A. Montanari, and S. Oh. Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 56(6):2980–2998, June 2010.

[11] S. Laky, P. Zaletnyik, and C. Toth. Compressing lidar waveform data. In *Proceedings of the International LIDAR Mapping Forum*, 2010.

[12] C. Mallet. *Using Full-Waveform LIDAR data for mapping in Urban Areas*. PhD thesis, Telecom Paris Tech, 2010.

[13] C. Mallet and F. Bretar. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1):1–16, January 2008.

[14] M. Rudelson and R. Vershynin. Sampling from large matrices: An approach through geometric functional analysis. *Journal of the ACM (JACM)*, 54(4), July 2007.

[15] A. Wehr and U. Lohr. Airborne laser scanning an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54:68–82, 1999.