REMOTE-SENSED LIDAR USING RANDOM SAMPLING AND SPARSE RECONSTRUCTION

Juan Enrique Castorera Martinez
Faculty Adviser: Dr. Charles Creusere
Klipsch School of Electrical and Computer Engineering
New Mexico State University

ABSTRACT

In this paper, we propose a new, low complexity approach for the design of laser radar (LIDAR) systems for use in applications in which the system is wirelessly transmitting its data from a remote location back to a command center for reconstruction and viewing. Specifically, the proposed system collects random samples in different portions of the scene, and the density of sampling is controlled by the local scene complexity. The range samples are transmitted as they are acquired through a wireless communications link to a command center and a constrained absolute-error optimization procedure of the type commonly used for compressive sensing/sampling is applied. The key difficulty in the proposed approach is estimating the local scene complexity without densely sampling the scene and thus increasing the complexity of the LIDAR front end. We show here using simulated data that the complexity of the scene can be accurately estimated from the return pulse shape using a finite moments approach. Furthermore, we find that such complexity estimates correspond strongly to the surface reconstruction error that is achieved using the constrained optimization algorithm with a given number of samples.

KEYWORDS

Compressive Sensing, Sparse Signal Reconstruction, LIDAR data modeling, LIDAR data transmission

INTRODUCTION

In this paper, we study the feasibility of a new paradigm in remote LIDAR data collection and transmission based on compressive sensing principles and using sparse reconstruction techniques. Specifically, we do *not* apply the conventional scenario here in which that LIDAR range data is collected by sampling the scene uniformly using a raster scan and the data points are then all transmitted (possibly with compression) over a communications link to a remote user for viewing or further processing. Rather, we assume here that range samples are collected randomly across subregions of the scene and that these range samples are again transmitted to the remote user where they are reconstructed for view or further processing. In this case,

however, sparse reconstruction techniques based on constrained l_1-norm optimization must be applied. The theory of compressive sensing or sampling tells us that reconstruction to a specified level of quality is possible with only a subset of randomly-selected range samples as long as the target scene is in some sense sparse, the details of which we discuss later. The major challenge here is to determine when enough LIDAR range samples have been collected so that we achieve the required reconstruction quality at the receiver (where the user of the data resides). While we could in theory run the reconstruction algorithm at the transmitter each time we acquire a new range sample, this is not practical due to the high computational complexity involved. Instead, we propose to analyze the temporal envelopes of the LIDAR return pulses in order to model the complexity of the surface being illuminated. Based on the estimated complexity, we can determine how many samples are required to achieve the desired surface accuracy.

In this paper, we consider the two stages of the process separately: surface complexity modeling based pulse complexities and the range sample / surface complexity correlation. In an actual system, one would tie these two processes together so that each time a LIDAR pulse was received, the surface complexity model would be refined as would also the estimated number of additional samples needed to achieve the reconstruction quality goal. Note as each pulse is received, the range value corresponding to that pulse can be transmitted through the communication channel to the user and that the user need not wait for all of the range samples before begin reconstruction of the target scene.

In the text that follows, we first introduce the basic theory of compressive sensing and discuss issues involved in applying it to LIDAR. In the next section, we present simulation results showing that we can connect the number of LIDAR return pulses required to achieve a given surface reconstruction accuracy to the fractal surface complexity. After this, we show that it is possible to estimate the surface complexity from the shape of the temporal envelopes of individual LIDAR return pulses. Finally, we present our concluding remarks and discuss future directions for this research effort.

COMPRESSIVE SENSING/SAMPLING THEORY

In 2006, the works of Candes, *et al.* and Donoho put forth the idea of compressive sensing, reminding those in the signal processing community that the often-quoted Nyquist sampling theorem provides only a sufficient condition for perfect signal reconstruction and not a necessary condition [2, 3, 4]. In particular, it was noted that if the underlying signal of interest can be viewed as being sparse with respect to some underlying basis (not necessarily the one in which its waveform would be most naturally sampled), then it is possible to sample that signal at a rate much lower than the Nyquist rate and still reconstruct it to some bounded precision. For example, natural images are sparsely represented in the wavelet domain and thus compressive sampling has been successfully applied to them [5].

In a compressive sensing-based system, we do not directly sense (sample) the continuous waveform *f(t)* containing the signal of interest. Instead, we sample the inner product of that waveform with some set of sampling functions: i.e.,

$$y_k = \langle f, \phi_k \rangle, \qquad k = 1, \ldots m \tag{1}$$

where ϕ_k is the *k*th sampling function and y_k is a sample of the signal. In an ideal system using

conventional sampling, $\phi_k(t) = \delta(t - kT)$ where T is the sampling period and $\delta(\bullet)$ is the dirac delta function while for non-ideal conventional sampling, the ϕ_k are shifted, nonoverlapping box functions [1]. Given an orthonormal basis $\Psi = \{\psi_1, \psi_2, \ldots, \psi_n\}$, we can write the expansion of signal *f(t)* in terms of this basis as

$$f(t) = \sum_{i=1}^{n} x_i \psi_i(t) \tag{2}$$

where $x_i = \langle f, \psi_i \rangle$. If *f(t)* is sparse with respect to this basis expansion, then only a small number of transform coefficients x_i need be retained to get a high quality reconstruction of *f(t)*.

From this point forward in our discussion, we will assume that our signal *f* is discrete and, specifically, that it is represented by a vector $\mathbf{f} \in \Re^n$. A discrete representation will not necessarily limit the utility of compressed sensing in LIDAR applications. For the discrete problem, (2) can thus be written as

$$\mathbf{f} = \Psi \mathbf{x} \tag{3}$$

where $\mathbf{\Psi}$ is an $n \times n$ matrix having as its columns the basis set $\{\psi_1, \psi_2, \cdots, \psi_n\}$ and $\mathbf{x} \in \Re^n$ contains our transform coefficients. The goal of compressed sensing in a discrete framework thus becomes that of representing **f** with *m* coefficients where $m << n$.

For compressive sampling to be effective, the sampling functions ϕ_k must be incoherent with respect to the basis function ψ_i. From [1], coherence between the sensing basis Φ and the representation basis Ψ is given by

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1 \le k, j \le n} \left| \langle \varphi_k, \psi_j \rangle \right|. \tag{4}$$

Thus, the coherence is defined as the magnitude of the largest correlation between any two elements of Φ and Ψ. If the basis in which our signal **f** is sparse is known *a priori*, then we can construct a sensing basis Φ which is maximally incoherent and thus reconstruct **f** with the best possible fidelity for a given number of measurements. Unfortunately, it is generally not possible to sense the signal using the optimal basis in most applications. Fortunately, it turns out that a random orthonormal sensing basis Φ will achieve with high probability a coherence of approximately $\sqrt{2 \log n}$ relative to any fixed sparse representation basis Ψ [1]. This value is considerably below the upper bound of $\sqrt{n}$ on the coherence in (4), and it justifies the widespread use of random sampling functions in compressed sensing.

Some interesting issues arise in the application of compressive sensing to the 'cloud of points' LIDAR problem. First, the points are not sampled on a uniform grid, making the direct application of the sensing equation (1) difficult without interpolation. Thus, our nonuniform set of range samples can be viewed as being randomly sampled with respect to an underlying uniform grid, effectively folding the interpolation process into the reconstruction optimization. A much larger problem here, however, is in the reconstruction process itself. Specifically, the convex optimization process is formulated as the following constrained l_0-norm minimization:

$$\min_{\tilde{\mathbf{x}} \in \mathbf{R}^n} \|\tilde{\mathbf{x}}\|_{l_1} \text{ subject to } y_k = \langle \phi_k, \Psi \tilde{\mathbf{x}} \rangle \quad \min_{\tilde{\mathbf{x}} \in \mathbf{R}^n} \|\tilde{\mathbf{x}}\|_{l_0} \text{ subject to } y_k = \langle \phi_k, \Psi \tilde{\mathbf{x}} \rangle \tag{5}$$

In words, this equation states that we are looking for the vector $\tilde{\mathbf{x}}$ having the fewest non-zero

elements (i.e., the maximum sparsity) that is consistent with the observations y_k. The reconstructed signal is then given by $\tilde{\mathbf{f}} = \Psi\tilde{\mathbf{x}}$. While this optimization problem can be solved using straightforward linear programming if approximated as an l_1-norm minimization, its solution requires implicitly that we know the transform basis ψ in which our signal of interest is sparse. For many well-studied applications such as image compression, this is not a major issue since basis for which the signal of interest is sparse is well known (e.g., a wavelet basis). In the case of LIDAR range maps, however, a general basis for which all maps are provably sparse is not known. Furthermore, we would contend that such a general basis does not exist because of the dependence of the received signal on not just sampling density but also on geometry, laser spot size, and the local complexity of the terrain. It is this fact that motivates the proposed compression algorithm.

Specifically, the algorithm that we are proposing analyzes and classifies each received range pulse at the encoder and uses this classification to determine the manner in which it should be encoded. This classification process is performed at the receiver and is ultimately used to model the complexity of subregions of the range map. From this complexity model we then extract the sparse transform bases required in the reconstruction process as well as an estimate of the number of range samples needed in the region to achieve, with high probability, acceptable reconstruction (e.g., applying Eq. (6) in [1]). In this paper, we consider the local range maps to be modeled by a collection of piecewise constant planes, and the mean plane size associated with the best fit of this model to the actual data is used to characterize subregions.

MODELING LIDAR SCENE SPARSITY USING COMPRESSIVE SENSING

In this section, we present simulation results showing that scene complexity is directly correlated to the number of locations that are randomly sampled by LIDAR pulses. The results in this section were originally presented in [6]. Random surfaces are generated by a fractal-based iterative algorithm which uses a midpoint displacement in two dimensions to create corner points delimiting smaller but geometrically equal shapes (e.g., square facets). The implemented fractal algorithm is the Diamond-Square (D-S) algorithm developed in [7]. This algorithm is implemented in combination with bilinear interpolation to change scene complexity. To reduce complexity, bilinear interpolation is used to fill the gaps between facets already defined by the fractal algorithm or by interpolation itself. In general, gaps are filled with facets linearly correlated with its surrounding neighbor facets, thus generating surfaces of lesser detail and randomness. The greater the number of facets generated with bilinear interpolation, the lower the surface complexity. In general, a surface of higher complexity will tend to be more random and more detailed than one of lower complexity but both will have similar shapes. Surfaces of lower complexity contain facets that have a more constant slope orientation. In contrast, surfaces of higher complexity contain a large number of small sized facets with random orientations, introducing higher detail into the surface. A total of seven surfaces of increasing complexity (i.e., facet sizes) which vary between 1 and 7 were generated.

For the laser pulse simulation, we assumed noiseless sampling and we set the laser footprint size equal to the size of the smallest facet in the surface of highest complexity. In other words, LIDAR signal returns with a surface of highest complexity will contain reflections containing the

information of a single intercepted facet. The power reflected by each facet and collected at the sensor site depends in this simulation only on the orientation of the facet's normal with respect to the pointing laser. The generated surface is randomly sampled with single pulses; lossless transmission and Lambertian-scattering upon reflection were further assumed.

Each surface is compressively sampled using a set of *m* random measurements defined by the sampling functions φ_k. In this study, the set of *m*-random measurements at which surface reconstructions are obtained was $m = \{4, 40, 400, 4000, 8000, 16000\}$. The reconstruction of the surface from an incomplete set of measurements using compressive sensing was achieved with the l_1 minimization program described in (5) using additional constraints. Estimates of the surface sparsity as a function of complexity were computed by establishing a relationship between the number of *m*-measurements required and the lowest achievable MSE using the available set of reconstructions. To test our approach, a total of a 1000 reconstructions were obtained for each of the *m*-measurements and generated surfaces of varying complexities. Combining these, results in a total of 42,000 reconstructions used for the estimation of the corresponding MSEs for each *m*-measurement and surface complexity. The reconstruction algorithm used is the total variation (TV) obtained from the l_1-magic Matlab collection of subroutines developed by Candés and Romberg in 2005 [8].

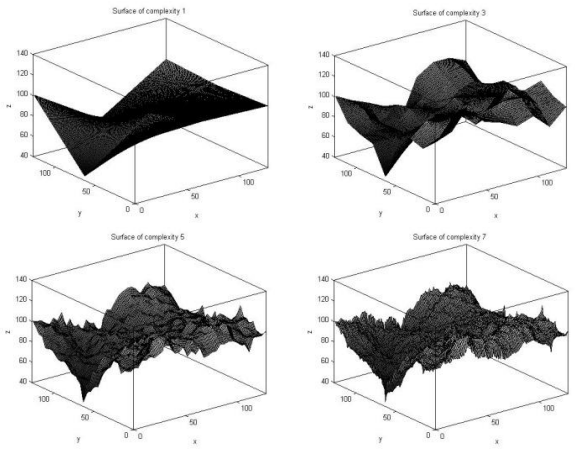

Figure 1: Generated surfaces a) Complexity 1, b) Complexity 3, c) Complexity 5, d) Complexity 7.

Each of the seven synthetic surfaces of size 129 × 129 was generated with a 100 meter mean level, a roughness of 30, and surface complexities ranging from 1 to 7. Examples of various surfaces are shown in Figure 1. In general, the generated surfaces have a similar shape but distinct levels of complexity. Surfaces of lower complexity contain less detail since the facets composing it have a correlated facet orientation among them while surfaces of higher complexity contain facets of increasing random orientations, thus introducing more detail. An example of the results of the compressive sensing reconstruction of a generated surface of complexity 4 using $m = 4000$ random measurements is shown in Figure 2. We note from figure 2 that the shape of the reconstructed signal resembles that of the original. The algorithm was not capable of recovering sharp edges formed between adjacent facets, however. The resulting mean squared error (MSE) of this reconstruction is of 5.713 with an approximate compression ratio of 4:1. To illustrate the resulting MSEs of all the reconstructions, the mean MSE was computed over the 1000 computed reconstructions for each value of *m* measurements and each surface complexity.

Values of *m* equal to 4, 40, 400, 4000, 8000, and 16000 were used. The results are plotted in Figure 3 which also includes plots of the mean MSEs of surfaces with distinct complexity. Note that the MSEs of the surface complexities are very close to one another for all sets of *m* random measurements, excepting that of the lowest complexity surface [6].

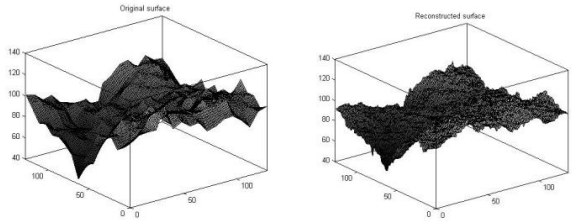

Figure 2: Surface reconstruction a) Original surface, b) Reconstructed surface

The results show that accurate reconstructions of the generated surfaces can be obtained using compressive sensing. Furthermore, Figure 3 shows that the number of measurements required to obtain small MSEs appears to increase as the complexity of the surfaces increase. For each of the surface complexities, the mean MSE for each *m* number of measurements was compared with the mean MSE corresponding to *m* = 16000. The minimum number of measurements for which the mean MSE for a given surface complexity is statistically equal to that for 16000 measurements was then selected as the minimum *m* satisfying Equation (6) in [1]. The resulting minimum *m's* corresponding to each of the surface complexities is shown in Table 1.

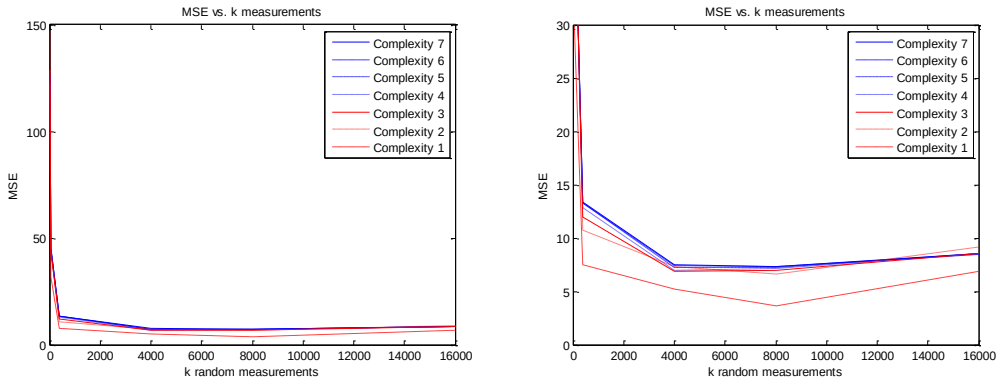

Figure 3: Reconstruction MSE as a function of k measurements. a) Mean MSE across reconstructions b) Zoomed plot

Table 1: Single return surface characterizations

Surface Complexity	Minimum *m*	Sparsity ratio *Sx* /*S*7
1	1232	0.375
2	1924	0.586
3	2769	0.844
4	2919	0.889
5	3127	0.953
6	3267	0.995
7	3282	1

The third column in Table 1 expresses a sparsity ratio relative to the surface of highest complexity as given by

$$\frac{m_i}{m_7} \geq \frac{C \cdot \mu^2(\Phi,\Psi) \cdot S_i \cdot \log n}{C \cdot \mu^2(\Phi,\Psi) \cdot S_7 \cdot \log n} \quad \text{(see [6] for details).} \qquad (6)$$

In (6), m_i is the number of measurements y_m required for a surface of sparsity i = 1 to 6 while m_7 is the number of measurements required a surface of complexity 7 (the highest). All of the terms on the right side of this equation cancel out except S_i / S_7 which are sparsity terms in the underlying theoretical formulation of compressive sensing [1]. Thus, this ratio directly relates the fractal surface complexity to the theoretical sparsity. From Table 1, we clearly see that as the complexity of the synthetically generated fractal surface increases, so too does the sparsity ratio and the corresponding number of samples required reconstruct the surface.

SURFACE COMPLEXITY ESTIMATION BASED ON LIDAR RETURN PULSE ENVELOP

Having shown that there is a direct correlation between the complexity parameter used to generate the random fractal surfaces and the required sampling density, we now investigate whether or not it is possible to estimate the complexity of the surface by analyzing LIDAR pulse returns. In particular, we evaluate the complexity of the temporal envelopes of the received LIDAR pulses by using the concept of finite rate of innovations and show that knowledge of these envelop complexities allows us to infer the complexity of the underlying surface and thus (based on the analysis of the previous section) the number of measurements required to achieve accurate sparse reconstruction. Full details of the proposed approach can be found where it was first presented in [9], so we will only summarize the approach and the results here. Random fractal surfaces are again generated exactly as before and Monte Carlo trials are applied to generate the statistical results.

A. Model of LIDAR Return pulse

Ideally, a single return waveform representing the topography of the small surface region illuminated by a single pulse (i.e. without scan) should be of the form of the weighted sequence of Diracs

$$x(t) = \sum_{k=0}^{K-1} a_k \delta(t - t_k), \quad t \in \Re \qquad (7)$$

where t represents continuous time, K the number of distinct illuminated facets (i.e., innovations) within the laser footprint and t_k indicates the TOF of the pulse caused by the k^{th} facet reflection. Weights a_k indicate the attenuation of power along the intervening path of the reflected pulse from the particular location corresponding to reflection k. The amount of power reflected depends on several factors, including the facet orientation with respect to the laser pulse, the material and the density. Because the laser pulse duration is not Dirac in practice, the return waveform itself will be composed of shifted and spread versions of this laser pulse. In this study, however, we model the system response using the *m-th* order *b*-spline function $\phi(t)$ (see [9] for details).

TABLE 2
CONFUSION MATRIX: SURFACE COMPLEXITY CLASSIFICATION

		TRUTH						
	Surface Complexity	1	2	3	4	5	6	7
PREDICTED	1	20	0	0	0	0	0	0
	2	0	20	0	0	0	0	0
	3	0	0	17	0	0	0	0
	4	0	0	3	16	1	0	0
	5	0	0	0	2	11	4	3
	6	0	0	0	2	6	11	15
	7	0	0	0	0	2	5	2

B. Constructing Model of LIDAR Pulses

Reconstruction of the signal parameters in (7) is achieved by sampling moments and using the finite rate of innovation (FRI) approach described in [10]. This approach reconstructs signals by sampling at their rate of innovation, a rate often much lower than the Shannon/Nyquist rate. In the case of *b*-spline reproducing polynomials, a choice of $N \geq 2K - 1$ samples suffices to model signal *x(t)* perfectly. In general, the procedure consists of first calculating the coefficients $c_{m,n}$ for which the following polynomial is satisfied

$$\sum_{n \in \Im} c_{m,n} \phi\left(t/T - n\right) = t^m, \quad m = 0,1,\mathsf{K}\ ,N \tag{8}$$

In addition, *b*-spline kernels are required to be of at least the same order as the degree of the reproducing polynomial for accurate signal reconstructions. In our case, coefficients corresponding to *b*-spline of order $N \geq 2K$ are required for retrieving the $2K$ signal parameters. These coefficients are then used to compute the signal moments for estimation of the signal parameters in (7) using the equality given by

$$\tau_m = \sum_n c_{m,n} y_n = \sum_{k=0}^{K-1} a_k t_k^m, \quad m = 0,1,\mathsf{K}\ ,N \tag{9}$$

This system of equations can be solved in matrix form for the time-of-flights (TOFs) t_k using a Yule-Walker system of equations involving at least $2K$ consecutive values of τ_m and then finding the roots of the polynomial (i.e., the annihilating filter) formed by the coefficients of the solution. The weights a_k can be found given the TOF locations $t_0, t_1, \mathsf{K}\ , t_k$ and arranging them in matrix form as a Vandermonde system following (5). A more detailed description of the procedure can be found in [10].

In addition to classifying return waveform complexity, we also build a simple classifier to determine the complexity of the surface under observation. Here, surface complexity can indicate both facet orientation randomness with respect to the pointing laser direction and the density of facets with similar orientation within the laser spot footprint. In general, we expect that reconstructed waveforms corresponding to surfaces of higher complexity contain fewer dependencies than those corresponding to lower complexities. Based on this, we train a classifier to determine surface complexity using the number of significant singular values necessary for reconstructing (in the least squares sense) the return waveforms at a pre-specified error

percentage.

C. Experimental Results

In our experiments here, we consider two different definitions of the underlying surface complexity: the complexity as given by the surface complexity parameter used to generate the fractal surface and the complexity as given by the size of the laser spot illuminating the target relative to the facet size. We term this latter quantity to be the laser footprint size (FPS). Table 2 shows a confusion matrix illustrating the accuracy of the pulse-based surface complexity analysis with a fixed FPS of 32x32 target facets and varying fractal complexity. A nearest neighbor classifier is used with a training set of 50 samples per class and an independent test set of 20 samples. Overall, we see that the results are very good for all of the classes except 7 (the highest complexity surface) which is often confused with class 6. Complexity estimation results based on footprint size are shown in Table 3. Here, the training set is composed of 5600 simulated LIDAR return pulses while the test set contains 1120 independently generated return pulses. The performance here is even better, ranging from 74% correct classification in the worst case to 99.5% in the best case.

TABLE 3
CONFUSION MATRIX: FPS CLASSIFICATION

			T	R	U	E
	fps (in facets)	2×2	4×4	8×8	16×16	32×32
PREDICTED	2 × 2	1115	248	0	1	0
	4 × 4	5	836	88	16	1
	8 × 8	0	36	998	116	22
	16 × 16	0	0	34	840	269
	32 × 32	0	0	0	147	828

CONCLUSIONS

In this paper we have studied the application of compressive sensing to LIDAR data acquisition. Specifically, we have simulated a target surface that is randomly generated, fractal, and piecewise-planar, assuming that each plane scatters light according to the typical Lambertian model. We then showed through Monte Carlo simulations that the number of randomly sampled LIDAR pulses used in a sparse reconstruction process was directly correlated to the fractal surface complexity and further that this complexity could be accurately predicted from the complexity of the received waveform as indicated by the innovations rate of its temporal envelop. We also showed that the characteristics of the temporal envelop are even more effective for predicting the number of individual facets illuminated by the laser spot—a quantity that can be viewed as another measure of scene complexity.

From the standpoint of telemetering, these results are useful because they open up the possibility of efficiently transmitting LIDAR range values from a remote platform back to a command center for reconstruction. We note that while the temporal envelops of the received pulses are processed at the LIDAR to model the scene complexity and thus determine whether or not enough samples have been collected, only the numeric range sample values need to be transmitted back to the command center. This combined with the fact that only a minimal

number of randomly-selected range samples are selected for each subregion in the scene could result in significant reductions in the required communications bandwidth over a conventional raster-scanned LIDAR system for which all of the range samples are transmitted regardless of the underlying scene sparsity. Furthermore, while the entire LIDAR return pulse envelop must be processed for complexity estimation in the LIDAR transmitter, the complexity of this processing is many orders of magnitude less than the alternative: a complete sparse reconstruction of the target surface each time a new return pulse is received.

Some areas of future work which we are interested in pursuing including improving our pulse complexity modeling by adding polarimetric signature analysis as well as collecting laboratory data with carefully ground-truthed target objects to validate these simulation results. In addition, we plan to study how a communication channel affects the reconstruction quality by examining the effects of quantization and channel errors.

REFERENCES

[1] E. Candes and M.B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21-30, March 2008.

[2] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[3] E. Candes and T. Tao, "Near optimal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.

[4] D. Donoho and X. Huo, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006.

[5] J. Romberg, "Imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 14-20, March 2008.

[6] Castorena, J.; Creusere, C.D.; Voelz, D.; , "Modeling lidar scene sparsity using compressive sensing," *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International* , vol., no., pp.2186-2189, 25-30 July 2010

[7] A. Fournier, D. Fussell, and L. Carpenter, "Computer Rendering of Stochastic Models", *Communications of the ACM*, June 1982

[8] "l1-magic", www.l1-magic.org.

[9] Castorena, J.; Creusere, C.D.; Voelz, D.; , "Using finite moment rate of innovation for LIDAR waveform complexity estimation," *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on* , vol., no., pp.608-612, 7-10 Nov. 2010.

[10] P.L. Dragotti, M. Vetterly and T. Blu. "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix", *IEEE Transactions on signal processing*, Vol.55, No. 5. May 2007.