MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

High-Resolution Lidar Using Random Demodulation

Boufounos, P.T.

TR2018-138 September 26, 2018

Abstract

Recently emerging applications, such as autonomous navigation, mapping, and home entertainment, have increased the demand for inexpensive and high quality depth sensing. In this paper we fundamentally re-examine the problem, considering recent advances in photoelectric devices, increased availability of fast electronics, reduced computation cost, and developments in sensing theory. Our main contribution is a real-time hardware architecture for time-of-flight (ToF) depth sensors that exploits random modulation to significantly reduce the acquisition burden. The proposed design is able to acquire compressive, critical, or redundant measurements, without requiring any hardware modifications, at the expense of small reduction in the system frame rate. The architecture we propose is sufficiently flexible to be operable in a variety of conditions and with a variety of reconstruction algorithms.

IEEE International Conference on Image Processing (ICIP)

HIGH-RESOLUTION LIDAR USING RANDOM DEMODULATION

Petros T. Boufounos

Mitsubishi Electric Research Laboratories
petrosb@merl.com

ABSTRACT

Recently emerging applications, such as autonomous navigation, mapping, and home entertainment, have increased the demand for inexpensive and high quality depth sensing. In this paper we fundamentally re-examine the problem, considering recent advances in photoelectric devices, increased availability of fast electronics, reduced computation cost, and developments in sensing theory.

Our main contribution is a real-time hardware architecture for time-of-flight (ToF) depth sensors that exploits random modulation to significantly reduce the acquisition burden. The proposed design is able to acquire compressive, critical, or redundant measurements, without requiring any hardware modifications, at the expense of small reduction in the system frame rate. The architecture we propose is sufficiently flexible to be operable in a variety of conditions and with a variety of reconstruction algorithms.

Index Terms— Depth sensing, random demodulation, compressed sensing, compressive LIDAR

1. INTRODUCTION

Recent advances in signal processing theory and acquisition hardware, such as compressive sensing (CS), have enabled significant improvements and cost reduction in many imaging applications. However, while these advances have resulted in noteworthy developments in a number of areas, the impact in time-of-flight (ToF) imaging has only been modest [1, 2]. Fortunately, thanks to progress in device technology, it is now possible for sensors and electronics to manipulate optical signals in a manner similar to other modalities. The only remaining hurdle in implementing such systems is the cost of fast signal acquisition.

In this paper we propose a new ToF architecture that uses an array of fast optical sensors, combined with individual random modulators that multiply each received signal. The modulated signals are added and a single fast analog-to-digital converter (ADC) is used to acquire their sum, instead of requiring one ADC for each sensor. Since the ADC is one of the most expensive components in ToF systems, the proposed architecture significantly reduces system cost.

The proposed architecture, inspired by [3], exploits recent advances in high-speed photonics and electronics. In particular, demand for high frequency optical and wireless communications has made the cost of very fast analog switches and multipliers negligible, and significantly decreased the cost of fast photodiodes and photodiode arrays. Thus, it is possible to manufacture the proposed architecture inexpensively, bar for the cost of the ADC. As we also demonstrate in this paper, advances in CS theory and practice further guarantee that this architecture is able to acquire and reconstruct the ToF of a properly modulated optical pulse with very high accuracy.

1.1. Contributions

The main realization in this paper is that the electronic components required to implement fast modulators are significantly cheaper and more flexible than the components required to implement optical light modulation or multi-channel ADCs. Thus, we contribute a flexible versatile hardware architecture for ToF imaging that

1. only requires a single high-speed ADC to acquire the signal,
2. allows for single-shot scene acquisition, enabling acquisition of scenes with motion,
3. enables sufficiently flexible operation with adjustable number of measurements, that can be selected at run-time, allowing for compressive, critical, and overcomplete acquisition,
4. provides wide system design versatility and can be easily incorporated in other designs, such as scanning systems and coded aperture systems.

1.2. Outline

The next section provides an overview of recently emerged compressive ToF architectures, as well as most common reconstruction approaches. Section 3 describes the proposed system architecture. Section 4 presents the system model we use for reconstruction and analysis. Section 5 describes how the acquired scene can be reconstructed, under different modeling assumptions and system parameters, as well as the computational trade-offs in the design. Section 6 concludes with simulations that validate the framework.

2. BACKGROUND

2.1. Compressive ToF Architectures

Earlier work on compressive ToF imaging relies mostly on adjustable coded aperture technology. In particular, several architectures [4, 5, 6] exploit digital micromirror devices (DMD) or other spatial light modulators (SLM) to create a coded aperture that can be modified during acquisition. The depth of a scene can then be acquired using a single detector and ADC by transmitting multiple pulses and taking snapshots, each acquired using a different coded aperture configuration. The hardware architecture is similar to the Rice single-pixel camera [7], using a time-resolved photodector and reconstruction algorithms specific to ToF imaging.

One drawback of adjustable coded aperture systems is that current DMD and other SLM hardware operates too slow for real-time scene acquisition, especially if the scene includes motion. Thus, a non-trivial time interval—typically in the order of hundreds of milliseconds—is required to acquire the necessary number of snapshots to reconstruct the scene. The reconstruction algorithms assume that the scene is static throughout the measurement process, which effectively precludes all but very slow-moving or static scenes.

Another drawback of coded aperture techniques is that approximately half of the reflected light arriving at the sensor is rejected by the code on the aperture. This incurs a 3dB signal-to-noise ratio (SNR) penalty at the sensor, which, in some applications can be significant, if not detrimental. Noise amplification is known to be an Achilles heel of compressive acquisition; the reconstruction SNR penalty is typically worse than the 3dB SNR loss at the sensor [8, 9].

More recently, it has been shown that a single, fixed, coded aperture can be used instead of an adjustable one, by increasing the number of sensors and ADCs [10, 11]. This enables depth acquisition using a single snapshot. In other words, it is possible to acquire scenes that include much faster motion, compared to adjustable coded aperture systems. Still, while the cost of a fixed coded aperture is trivial compared to the cost of the system, the cost trade-off is only advantageous if the additional detectors and, especially, ADCs are less expensive that a DMD or other SLM device.

Unfortunately, a fixed coded aperture also induces the 3dB SNR penalty at the receiver. The single-shot ability allows for pulses of longer duration in order to introduce more energy into the scene and, thus, improve the SNR, without incurring significant motion blur. However, it would still be desirable to capture all the light arriving at the receiver and use a shorter pulse.

The proposed architecture bears some similarities with earlier approaches. On one hand, similarly to the architectures using an adjustable coded aperture, we only require a single ADC. On the other hand, similarly to the architectures using a fixed coded aperture, we require several detectors. Fortunately, fast photodiodes with sub-nanosecond response times are widely available and inexpensive, especially when implemented in an integrated circuit.

There are also significant differences. In particular, our architecture does not require a coded aperture, thus capturing all of the incident light. While the function of the electronic modulation can, in some sense, be interpreted as an electronically adjustable coded aperture, the behavior is quite different, allowing for single shot imaging. An alternative interpretation is that the modulation code is applied in time domain whereas coded apertures in earlier efforts implement a spatial domain code. Still, the proposed architecture can be combined with a fixed optical coded aperture, to reduce the required number of sensors. While this is a straightforward combination with [11], we do not pursue it in this paper.

2.2. Compressive Sensing and Scene Reconstruction

The architectures described above rely on the emergence of CS and sparse regularization as a powerful signal reconstruction framework [12, 13, 14]. As well established by now, CS exploits sparse signal priors to significantly reduce the number of measurements required to acquire a signal that fits the assumed model. By regularizing the reconstruction with the appropriate model, it is possible to provide reconstruction performance and stability guarantees.

The most common signal priors are sparsity in some basis, commonly enforced using ℓ_1 regularization or greedy algorithms, or sparsity in the signal gradient, commonly enforced using Total Variation (TV) regularization [14]. Signals with finite rate of innovation (FRI) are a continuous-time model alternative, more appropriate for capturing short pulses with unknown timing [15]. More sophisticated models can be enforced using model-based compressive sensing [16]. Frequently, the reconstruction algorithm is specifically designed and tuned in tandem with the hardware acquisition system.

Particularly for depth sensing, the coded aperture-based architecture in [5, 6] relies on a two-step acquisition approach: an FRI model is first used to detect the depths at which reflectors are located;

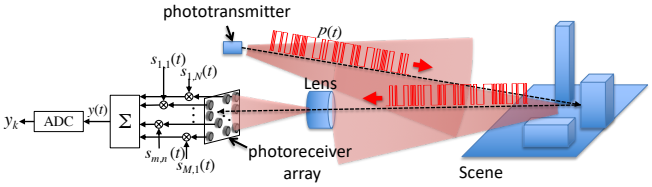

Fig. 1. Overview of the proposed system architecture and operation

the reflectivity is then recovered separately at each depth, optionally using sparse regularization. Alternatively, [10, 11] demonstrate that a 3-dimensional (3D) representation of the scene allows for linear modeling of the acquisition, thus enabling one-step sparsity-based recovery methods. Furthermore, application-specific models can improve reconstruction quality. However, TV regularization of the depth map is not straightforward in this representation. Computation becomes a significant bottleneck, even more so than in conventional sparse recovery approaches. Still, the model is more general and can be used in a number of systems, including the ones in [4, 5, 6]. For this reason, we adopt the same 3D representation in this paper.

3. SYSTEM ARCHITECTURE

A schematic of the system architecture we propose, and its operation, is shown in Fig. 1. The system comprises of a single fast phototransmitter, such as a light emitting diode (LED) or a laser, transmitting a coded pulse to the scene to be measured. The light from the code pulse floods the scene and illuminates all objects. It arrives at each object with different delay, depending on their distance from the source. The code is applied in time by rapidly turning the transmitter on and off.

The light incident in the scene is reflected from the objects and received by an array of photodetectors through a focusing system. Each photodetector corresponds to a pixel of the depth image. The light reflected back from the scene is delayed according to the distance of each object in the scene from the corresponding pixels receiving its reflection. The signal received by each photodetector is modulated with a known randomized modulation, unique to the photodetector. Finally, all the modulated signals are added and their sum is acquired by an ADC.

The whole system, from transmission to reception, operates at a bandwidth commensurate with the time resolution required by the system. Thus, for practical purposes, the phototransmitter, the photodetectors and the modulators, need to have sub-nanonsecond response times to achieve round trip ToF resolution lower than 30cm, i.e., depth resolution of 15cm. This is equivalent to bandwidth of a few GHz. Phototransmitters, photodetectors and modulators are readily available at this bandwidth. The ADC dominates the cost of the system. We should note, that the time resolution can be improved using computational methods such as FRI modeling, at the expense of higher computational cost.

The architecture can be easily modified to either improve performance or reduce the number of components. If cost permits, multiple ADCs may be used in a variety of ways. For example, each signal from the photodetector could be modulated by multiple unique modulation sequences, each feeding to a different summation and ADC circuits running in parallel. Thus, each sensor is acquired by all ADCs. Alternatively, the array can be split in smaller groups of sensors, each group handled by a single summation and ADC circuit under the same modulation architecture described above.

We exploit temporal mixing and subsampling to reduce the ADC cost. However, if the reconstruction algorithm incorporates spatial regularization, spatial mixing and subsampling can be used to also reduce the number of sensors. For example, with the addition of a fixed coded aperture, sensors can be removed from the system, to obtain a subsampled array similar to [11].

4. MODELING AND OPERATION

4.1. System Model

The phototransmitter emits a pulse $p(t)$ in time, of duration T_{p}. For modeling convenience, although not necessary, we assume the pulse comprises of a sequence of K_{p} square pulses $p_o(t)$ of duration τ, modulated by a discrete sequence p_k of length $K_{\mathrm{p}} = T_{\mathrm{p}}/\tau$. We assume this ratio is integer.

$$p(t) = \sum_k p_k p_o(t - k\tau), \tag{1}$$

$$p_o(t) = \begin{cases} 1 & 0 \le t < \tau \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

In other words, p_k is a discrete representation of $p(t)$ under the basis comprising of shifts of $p_o(t)$. Note that $p(t)$, and, consequently p_k is positive, since the transmitter cannot transmit negative light. Although not necessary in the subsequent development, to simplify hardware implementation, we use $p_n \in \{0, 1\}$.

Any object in the scene at distance d will receive the pulse with delay d/c—where c is the speed of light—and reflect it attenuated, according to the object's reflectivity. The pulse reflection is further attenuated inversely proportional to the square of distance of the object and returns to the receiver with delay $2d/c$. In other words, given a delay t, the corresponding distance is $tc/2$.

The reflections are received by an array of $M \times N$ detectors, i.e., pixels, focused on the scene, which measure the time profile of the pulse, denoted $r_{m,n}(t)$. The received pulse is modulated by a unique coded sequence, $s_{m,n}(t) = \sum_k s_{m,n,k} p_o(t - k\tau)$, where $s_{m,n,k}$ is the discrete representation of the code. Since this is an electronic and not an optical modulator, this code can be negative. Even though not assumed in the subsequent development, to facilitate hardware implementation, we use $s_{m,n,k} \in \{\pm 1\}$ or $\{-1, 0, +1\}$. Thus, the multiplication can be implemented using a simple inverter and an analog switch. In many cases, an analog multiplier may be less expensive, providing more freedom in designing $s(t)$.

All modulated received pulses are summed using an adder

$$y(t) = \sum_{m,n} r_{m,n}(t) s_{m,n}(t). \tag{3}$$

The sum is sampled at rate τ to produce the measurements

$$y_k = \sum_{m,n} r_{m,n}(k\tau) s_{m,n}(k\tau) \tag{4}$$

$$= \sum_{m,n} r_{m,n,k} s_{m,n,k}, \tag{5}$$

where $r_{m,n,k} = r_{m,n}(k\tau)$ and $s_{m,n,k} = s_{m,n}(k\tau)$ are samples, at rate τ, of $r_{m,n}(t)$ and $s_{m,n}(t)$, respectively.

4.2. Scene Representation and Acquisition

To represent the scene, we use the same discretization. Specifically, we use m and n to index the pixel of the scene, i.e., the dimensions parallel to the plane of the sensors. We use k to index the depth of the scene, using distance resolution corresponding to a roundtrip travel of the pulse during the sampling period τ. In other words, at depth index k, the corresponding true depth is $d = k\tau c/2$. Under this discretization, we denote the reflectivity of each voxel in the scene using $x_{n,m,k}, (n, m, k) \in \{0, \dots, N\} \times \{0, \dots, M\} \times \{0, \dots, K_{\mathrm{s}}\}$, in which we include the attenuation due to distance.

Assuming the scene does not contain diffusive media, such as smoke or fog, along any spatial direction (n, m) we expect to observe very few depths k with non-zero reflectivity $x_{n,m,k}$. In particular, we may assume that the scene only contains opaque reflectors and that the discretization is sufficiently fine that every voxel is completely covered spatially if it contains a reflector. In this case, we expect at most a single non-zero along each direction (n, m). This is because, under these assumptions, a reflector at a certain distance will obscure any reflectors behind it and will only be visible if there is no reflector in front of it. Thus, if the reflector reflects the pulse, it means that the recoverable reflectivity in front of and behind this reflector should be zero. Even if we assume partial reflectors, e.g., semi-transparent surfaces, or reflectors that partially cover a voxel and allow some light to be transmitted, we still expect very few voxels of non-zero reflectivity along any direction (n, m).

Given this model, the received signal at each sensor is a convolution of the scene along the corresponding direction with the pulse $p(t)$. The equivalent discrete model is $r_{m,n,k} = x_{m,n,k} \otimes_k p_k$. In other words, the measurements are equal to

$$y_n = \sum_{m,n} ((x_{m,n,k} \otimes_k p_k) s_{m,n,k} \tag{6}$$

Using vector and operator notation, we denote the measurements as

$$Y = \Sigma(S(P(X))) = A(X), \tag{7}$$

where X denotes the scene, $P(\cdot)$ denotes the operator implied by the convolution with the pulse, $S(\cdot)$ denotes the diagonal operator implied by the modulation, $\Sigma(\cdot)$ denotes the summation operator, and $A(\cdot)$ denotes the whole measurement operator. Given the square pulse assumption in (2), it is straightforward to show that the location of the reflector within the voxel does not affect the discretization. Thus, under this assumption, the sampling rate τ, and the corresponding discretization fully determines the system's depth resolution.

Depending on the length of the pulse, compared to the length of the scene, the system is not necessarily compressive. Given a scene maximum length K_{s} and pulse length K_{p}, the measurements have the same length as the resulting convolution, i.e., length $K_{\mathrm{p}} + K_{\mathrm{s}} - 1$. If this is greater than the size of the scene, i.e., if

$$K_{\mathrm{p}} + K_{\mathrm{s}} - 1 \ge MNK_{\mathrm{s}} \tag{8}$$

$$\Leftrightarrow K_{\mathrm{p}} \ge (MN - 1)K_{\mathrm{s}} + 1, \tag{9}$$

then the number of measurements are greater than the size of the scene. In that case, the measurement operator (7) can be designed to be full rank.

To guarantee incoherence, we randomize the system through the selection of the light pulse sequence p_k and the modulation code $s_{m,n,k}$. Specifically, p_k is determined using an independently and identically distributed (i.i.d.) Bernoulli random variable, taking values in $\{0, 1\}$ with equal probability. Similarly, $s_{m,n,k}$ is drawn from an i.i.d. distribution taking values in $\{\pm 1\}$ with equal probability. Note that in [3], incoherence is guaranteed since the signal model assumes the signal is sparse in frequency. Since this is not the case here, a long and varying pulse $p(t)$ is necessary to guarantee incoherence with the time modulation and summation at the receiver.

5. RECONSTRUCTION APPROACHES

5.1. Adjoint Operator and Backprojection

Given the forward model (7), scene reconstruction becomes an inverse problem. Depending on the length of the pulse, different reconstruction algorithms can be used. However, any reconstruction algorithm requires the use of adjoint of the forward operator (7). This is straightforward to compute using the composition property of adjoint operators: For any two linear operators $A(\cdot)$ and $B(\cdot)$, $(AB)^* = B^*A^*$, where $(\cdot)^*$ denotes the adjoint. Thus, the adjoint of (7) can be simply computed as

$$A^*(Y) = P^*(S^*(\Sigma^*(Y))). \tag{10}$$

The adjoint of the summation Σ is a replication to all channels. The adjoint of the diagonal modulation operator S is also a modulation by the same sequences, i.e., $S^* = S$. Finally, the adjoint of the convolution by the pulse, P, is a convolution by the same pulse, time reversed, i.e., a matched filter. In summary, the adjoint computes the backprojection of the data, i.e., a scene $\widetilde{X}$ such that

$$\widetilde{x}_{m,n,k} = (y_k s_{m,n,k}) \otimes_k p_{-k}. \tag{11}$$

5.2. Scene Models, Reconstruction, and Complexity

As described in (9), if the pulse is short, the acquisition is compressive. In that case, it becomes necessary to use regularization to recover the scene. Depending on the pulse length and assumptions on the scene, stronger or lighter regularization might be appropriate. If the acquisition is more compressive, using short pulses, then model-based regularization techniques, such as the ones described in [10, 11] would be necessary. If the scene includes semi-transparent objects, model-based algorithms may fail; models in [10, 11] do not accept multiple reflections along a single direction (m, n). In that case, and using longer pulses, conventional sparsity models, using a convex reconstruction or a greedy algorithm would perform well. Adapt streaming reconstruction algorithms, such as [17], could also achieve faster reconstruction of time varying scenes, possibly with fewer measurements.

With a sufficiently long pulse, satisfying (9), the forward operator becomes full rank and least-squares inversion is sufficient. However, even in these cases, regularization or thresholding can improve performance by denoising the reconstruction.

As the transmitted randomized pulse becomes even longer, the operator becomes overcomplete, i.e., redundant. In this case, it is straightforward to show that, asymptotically, the forward operator becomes a tight frame. Thus, if the pulse is sufficiently long, simple backprojection using the adjoint operator (11), followed by thresholding approximately inverts and denoises the acquisition. We defer the proof to an extended version of the paper.

5.3. Reconstruction Complexity Trade-offs

Inversion by simple backprojection and thresholding is particularly appealing in practical applications, especially in real-time ones. The computational complexity of the backprojection operator is orders of magnitude lower than typical sparse regularization or, even, least squares algorithms. Typical sparse recovery algorithms require hundreds of iterations, where each iteration applies both the forward and the adjoint operator. Thus, each iteration exhibits at least twice the computational cost of a simple backprojection.

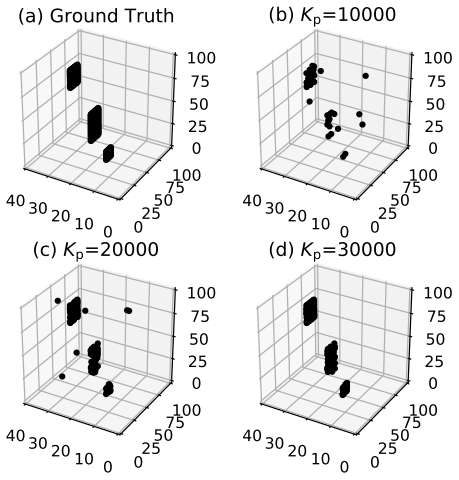

Fig. 2. Performance for different pulse lengths. (a) Original Scene, (b)–(c) Backprojected thresholded reconstruction for different K_p

The typical assumption in compressive sensing is that the acquisition cost per measurement is significant, compared to the computational cost of sparse reconstruction. However, in our architecture, the cost of a longer pulse is minimal. Specifically, the main cost is the additional energy used to transmit the pulse and the longer time required to acquire the scene—which implies a lower maximum frame rate. There is no requirement for additional sensors, or additional ADCs. The very high pulsing rate makes this cost insignificant, enabling very fast reconstruction by backprojection.

As a simple example, consider a system that images a 30m deep scene using $\tau = 200$ps, i.e., has 3cm depth resolution and $K_\mathrm{s} = 100$. Assume the system has $MN = 0.25$MP (megapixel) spatial resolution, e.g., $M = 500$, $N = 500$. Using (9), we obtain

$$K_\mathrm{p} \geq (MN - 1) * K_\mathrm{s} + 1 \approx MNK_\mathrm{s} \tag{12}$$

$$\implies K_\mathrm{p} \gtrsim 2.5 \times 10^7, \tag{13}$$

which is equivalent to a $T_\mathrm{p} = 5$ms pulse. Even if use a pulse length 5 times longer, i.e., 25ms, to improve the operator conditioning, we can still achieve 40 frames per second (fps) frame rate for the system.

6. EXPERIMENTS

To verify our approach, we conducted an experiment on a $M = 100$, $N = 100$ scene, with depth $K_\mathrm{s} = 40$ and three extended reflectors. The scene is acquired with pulses of different lengths, and reconstructed using backprojection and thresholding. Figure 2 plots (a) the original scene and (b)–(c) the backprojected thresholded reconstruction for $K_\mathrm{p} = 10000$, 20000, and 30000, respectively. As expected, increasing the pulse length improves reconstruction quality.

Note that the pulse length required to obtain non-compressive acquisition according to (9) is $K_\mathrm{p} \simeq NMK_\mathrm{s} = 400000$, i.e., at least one order of magnitude higher than the pulse lengths in the experiments. Still, even at this low pulse length, the randomization makes the acquisition operator sufficiently incoherent that backprojection followed by thresholding performs very well. Of course, we expect that more sophisticated reconstruction methods will perform even better with even shorter pulse lengths. However, as we argue above, the computational cost does not make the trade-off favorable.

7. REFERENCES

[1] Shahram Izadi, Ayush Bhandari, Achuta Kadambi, and Ramesh Raskar, "3D imaging with time of flight cameras: theory, algorithms and applications," in *ACM SIGGRAPH 2014 Courses*. ACM, 2014, p. 5.

[2] Achuta Kadambi, Ayush Bhandari, and Ramesh Raskar, "3D depth cameras in vision: Benefits and limitations of the hardware," in *Computer Vision and Machine Learning with RGB-D Sensors*, pp. 3–26. Springer, 2014.

[3] Joel A Tropp, Jason N Laska, Marco F Duarte, Justin K Romberg, and Richard G Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 520–544, 2010.

[4] Gregory Howland, Petros Zerom, Robert W Boyd, and John C Howell, "Compressive Sensing LIDAR for 3D imaging," in *CLEO: Science and Innovations*. Optical Society of America, 2011, p. CMG3.

[5] Ahmed Kirmani, Andrea Colaço, Franco N. C. Wong, and Vivek K. Goyal, "Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor," *Opt. Express*, vol. 19, no. 22, pp. 21485–21507, Oct 2011.

[6] Ahmed Kirmani, Andrea Colaço, Franco NC Wong, and Vivek K Goyal, "CoDAC: a compressive depth acquisition camera framework," in *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2012, pp. 5425–5428.

[7] Marco F. Duarte, Mark A. Davenport, Dharmpal Takhar, Jason .N. Laska, Ting Sun, Kevin F. Kelly, and Richard G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, March 2008.

[8] Ery Arias-Castro and Yonina C. Eldar, "Noise folding in compressed sensing," *IEEE Signal Processing Letters*, vol. 18, no. 8, pp. 478–481, Aug 2011.

[9] Mark A. Davenport, Jason N. Laska, John R. Treichler, and Richard G. Baraniuk, "The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4628–4642, Sept 2012.

[10] Petros T. Boufounos, "Depth sensing using active coherent illumination," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Kyoto, Japan, March 25-30 2012.

[11] Achuta Kadambi and Petros T. Boufounos, "Coded aperture compressive 3-d LIDAR," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Brisbane, Australia, April 19-24 2015.

[12] Emmanuel J Candes, Justin K Romberg, and Terence Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on pure and applied mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[13] David L Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[14] Simon Foucart and Holger Rauhut, *A mathematical introduction to compressive sensing*, vol. 1, Birkhäuser Basel, 2013.

[15] Martin Vetterli, Pina Marziliano, and Thierry Blu, "Sampling signals with finite rate of innovation," *IEEE transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, 2002.

[16] Richard G Baraniuk, Volkan Cevher, Marco F Duarte, and Chinmay Hegde, "Model-based compressive sensing," *IEEE Trans. Info. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[17] Petros T. Boufounos and M. Salman Asif, "Compressive sensing for streaming signals using the streaming greedy pursuit," in *Proc. Military Communications Conference (MILCOM)*, San Jose, CA, October 31 - November 3 2010.