STRUCTURED REPRESENTATION-BASED ROBUST AGILE-BEAM LADAR IMAGING

Vishal M. Patel\textsuperscript{1} and Michael A. Powers\textsuperscript{2}

\textsuperscript{1}Center for Automation Research, University of Maryland, College Park, MD 20742
\textsuperscript{2}General Dynamics, 1231 Tech Court, Westminster, MD 21157

pvishalm@umiacs.umd.edu, mpowers@gdrs.com

ABSTRACT

Laser Detection And Ranging (LADAR) imagers operate by measuring the distance to an object by timing how long a transmitted pulse takes to make a round trip between the transmitter and the receiver. LADAR images are often characterized by a multiplicative noise known as speckle which often makes the interpretation of data difficult. In this paper, we propose a more general formulation of the problem of reconstructing the piecewise smooth and texture component (speckle) directly from a set of LADAR measurements. The proposed method is evaluated on the real data collected by using a new agile-beam LADAR architecture.

Index Terms— Agile-beam LADAR imaging, image reconstruction, computational imaging.

1. INTRODUCTION

Laser Detection And Ranging (LADAR) is a technology that uses laser light to measure the distance to an object. The idea is that when light shines onto an object, a portion of that light is reflected back. The time between when the laser is fired and when the reflected light is received is used to measure the range of a target. Many LADAR systems use some form of beam scanning, such as a spot or line, to resolve transverse features of an object. If the laser is scanned over a region of interest, 3D image of the object can be reconstructed [1]. Other systems, often called Flash LADAR, use a relatively uniform beam to illuminate an object and an array of detectors to resolve transverse features. Our Agile-beam LADAR concept uses illumination modulation in time (i.e., a pulse along the direction of propagation) and space (i.e., a beam with patterning in the plane perpendicular to the direction of propagation) to measure a series of inner products so that they may reconstruct a three dimensional image of objects in a scene.

Our concept is intended for robotic perception to minimize size, weight, power and cost. Our initial effort explores the system design and computational imaging algorithms as we look forward toward nanophotonic optical phased array devices. In contrast to contemporary systems implemented in bulk optics, such optical phased arrays suggest future LADAR systems bearing greater similarity to today’s active electronically-scanned radar systems and enabling considerable advantages in moving operations that are presently implemented in optical hardware to software in digital electronic computers.

Recent progress in phased arrays at optical frequencies has been encouraging [2]. We expect that continued device development will attain high densities, rapid modulation, and high efficiencies needed to use a single array for transmission and reception in a monostatic configuration. The wavefront control to be afforded by this technology will go beyond beam steering to allow the formation of arbitrary patterns with frequency, phase, and amplitude agility to create scan patterns that adapt to changing environments and mission requirements. Moreover, the size and weight envelope of a phase-array-based LADAR is attractive, since the optics of the system suggest the form factor of a thin disk [3]. Our approach to suppressing speckle in reconstructed images is especially significant to imaging concepts using optical phased arrays. Practical implementations of these devices may (initially and perhaps persistently) only afford phase modulation which results in speckle intrinsic to the illumination pattern, in addition to speckle imparted by a rough scattering surface. The algorithm we present here may greatly simplify optical device construction and overcome physical limitations of wavefront generation.

The combination of arbitrary wavefront control, ample computational resources (the cornerstone of any robotic platform), and the nature of the desired output product suggests that this approach, and perhaps computational imaging in general, is remarkably well suited to robotic applications. In one sense, all robotic perception is computational by default, at least on the interpretation side where image fusion, stereo vision, and object recognition algorithms are used to build an output product: a semantic world model of the robots surroundings, which is a sparse representation compared to the input measurements. The Agile-beam LADAR concept envisions a computational approach on both the measurement apparatus, which currently is performed by the analog optical operation of lenses and focal planes, and the subsequent interpretation apparatus, which is performed on digital electronic computers. A seamless, software-based architecture melding measurement and interpretation suggests more efficient measurements that may be performed in a basis other than the pixel basis, which is otherwise rigidly defined in optical hardware. The precedent set by software-defined radio and cognitive radio is a useful analogy.

Compressive sensing (CS) implemented in the Rice single-pixel camera is one example of how measurement in a non-pixel basis yields both measurement efficiency and permits advantageous hardware architectures (i.e., a focal point receiver rather than a focal plane array) [4]. It is important to note that pixel-basis reconstruction considered essential in many computational imaging approaches is mostly superfluous in a robotic perception system. To the extent that objects and surfaces can be measured and identified in a compressed or alternative basis, as demonstrated by smashed-filter approaches, pixel basis reconstruction is unnecessary [5]. Moreover, the intelligence sub-systems of robotic vehicles are able to feed the perception-sub system with a set of priors based on the robots understanding of the environment and its situational awareness, thereby simplifying the identification process and allow tuning of the optical
parameters that would otherwise be impossible if they were rigidly defined in hardware.

With the Agile-beam concept, we take a prospective step toward phase array architectures using more efficient non-pixel basis sampling. In the case of active sensing with LADAR, compressive sensing techniques are especially beneficial since voxels are quite expensive owing to the laser power, signature emission, and avalanche photodiode arrays with sophisticated read-out electronics required to produce them. However, one of the assumptions made when applying CS theory for imaging is that the underlying image is sparse in some basis. This was shown to be a somewhat adequate assumption for LADAR imaging in [6], [7], [8]. However, LADAR images are often characterized by a multiplicative component known as speckle. Speckle can make the compressibility and interpretation of LADAR images more difficult. In this paper, we present an improved algorithm that can reconstruct LADAR images by recovering two components of a LADAR image directly from a set of measurements. These components represent piecewise smooth and textured (speckle) parts of a LADAR image.

The rest of the paper is organized as follows. In Section 2, we describe our agile-beam LADAR system architecture. The LADAR observation model and the optimization algorithms are described in Section 3. Experimental results are presented in Section 4 and Section 5 concludes the paper with a brief summary and discussion.

2. AGILE-BEAM LADAR ARCHITECTURE

The concept described in this paper was implemented using off-the-shelf components, including the “bucket detector” receiver used in the ARL MEMS-scanned LADAR [9]. A key compromise versus an ideal implementation is the use of a liquid-crystal-on-silicon (LCOS) spatial light modulator as a surrogate for a bona-fide optical phased array described in the introduction. While this liquid crystal SLM implements the essential function of an optical phased array, wavefront modulation/generation, it lacks amplitude control, requires an externally formed Gaussian beam for illumination, and is limited to relatively slow switching speeds (less than 1 kHz). Despite these shortcomings, it is an effective surrogate to demonstrate the concept.

The arrangement of components is shown in Figure 1. A linearly-polarized transmitter beam at 1550 nm, shown in red, is generated by a 2-nsec pulsed fiber laser and collimated through a multi-element lens (not shown). A mirror is positioned to align the beam to the center of the SLM and remains fixed in that position. The beam reflects from the SLM, passes through baffles intended to block stray light, and illuminates a distant target which is about 3 meters from the aperture. The backscatter is collected by the bucket detector receiver, which has a field of view that subtends the entire area projected by the SLM-generated far field patterns. The receiver feeds a high speed digitizer.

Three dimensional images are generated by recording time-domain inner product measurements between the three dimensional scene and a series of pulsed far-field illumination patterns whose time properties are defined by a laser pulse and whose cross-range spatial properties are defined by a computer-generated hologram written to the SLM. The system is configured to generate voxels in a \( n \times n \times 7 \) data cube, where \( n \) is the number of cross-range image elements in each of two spatial dimensions and \( T \) is the number of image elements in the direction of beam propagation (i.e., time). For the tests in this paper, a \( 32 \times 32 \times 301 \) image datacube was generated. The 301 elements in time are fully sampled and digitized as with any other full-waveform LADAR system, and each sample is one range “slice”. The \( 32 \times 32 \) cross-range elements may be measured in a variety of modes. The first mode, meant for comparison to classic techniques and to demonstrate flexibility, is raster scanning. In raster mode, 1024 wavefront-generating digital holograms are generated, each defining one beam deflection in the \( 32 \times 32 \) cross-range sampling grid. Naturally, the digital holograms required to affect these wavefronts have Fresnel prism-like 2D phase profiles. As an alternative to raster scanning using the identity matrix, we define 1024 digital holograms whose far-field illumination patterns are rows of the discrete cosine transform matrix and another mode with 1024 digital holograms that project patterns of the rows of the Hadamard transform matrix [10]. Each of the 1024 patterns in any mode generates a 301 sample time domain waveform, recording the backscatter range profile for a given illumination pattern. Inverting the measurement matrix transforms the measurements to the pixel basis.

Operating the prototype in a CS mode is similar to the other modes, except a random (rather than orthonormal) measurement matrix is used, and fewer than 1024 laser pulses are used to generate the full \( 32 \times 32 \times 301 \) image data cube.

3. OBSERVATION MODEL AND OPTIMIZATION

Assuming that the range of interest is known a priori, the measurement process can be written as

\[
y = Ax + b, \tag{1}
\]

where \( x \in \mathbb{R}^N \), \( N = n^2 \) is the cross-range element, \( A \in M \times N \) with \( M < N \) is the measurement matrix when CS used and \( b \) is the additive measurement noise. Furthermore, assuming that the range slice is a superposition of two different signals, (1) can be rewritten as

\[
y = A(x_p + x_t) + b = Ax_p + Ax_t + b, \tag{2}
\]

where \( x_p \) and \( x_t \) are the piecewise smooth component and texture component of \( x \), respectively. We further assume that \( x_p \) is sparse in an \( N \times N_p \) dictionary represented in matrix as \( D_p \), and similarly, \( x_t \) is sparse in an \( N \times N_t \) dictionary represented in matrix form as \( D_t \) so that \( x_p = D_p \alpha_p \) and \( x_t = D_t \alpha_t \) for the piecewise smooth and

---

1See [11], [12] for theoretical details on decomposing an image into the sum of a piecewise smooth part and a textural part.
the texture component, respectively. The sizes $N_p$ and $N_t$ are typically much larger than $N$. The texture dictionary $D_t$ should contain atoms that are oscillatory in nature such as those found in the discrete cosine/sine transform and the Gabor transform. The dictionary $D_p$ should be able to process images with geometric features such as edges. The matrix $D_p$ should be some type of wavelet, shearlet, curvelet, or contourlet dictionary.

Using this decomposition, the LADAR measurement process at a given range slice can be reformulated as

$$y = AD_p\alpha_p + AD_t\alpha_t + b = B_p\alpha_p + B_t\alpha_t + b,$$

where $B_p = AD_p$ and $B_t = AD_t$. We propose to recover the LADAR image $x$ by estimating the components $x_p$ and $x_t$ by solving the following problem:

$$\alpha_p, \alpha_t = \arg \min_{\alpha_p, \alpha_t} \lambda \|\alpha_p\|_1 + \lambda \|\alpha_t\|_1 + \gamma TV(D_p\alpha_p) + \frac{1}{2}\|y - B_p\alpha_p - B_t\alpha_t\|_2^2,$$  \hspace{1cm} (4)

where $\gamma$ and $\lambda$ are two regularization parameters and $TV$ is the total variation (i.e. sum of the absolute variations in the image). The two components are the corresponding representations of the two parts and can be obtained by $x_p = D_p\tilde{\alpha}_p$ and $x_t = D_t\tilde{\alpha}_t$. Once the two components of $x$ are recovered, we obtain the final estimate as

$$\hat{x} = \tilde{x}_p + \tilde{x}_t.$$  \hspace{1cm} (5)

This notion of separating a signal into different morphologies using sparse representations is often known as Morphological Component Analysis (MCA) [12].

3.1. Speckle

In LADAR imaging, the multiplicative speckle noise model is often given by $z = x \cdot v$, where $z$, $x$ and $v$ are the noisy LADAR image, the noise free image and speckle considered as a random vector, respectively. Here, $\cdot$ represents the point wise multiplication. See [13] for an extensive investigation of the statistics and the origins of speckle. This multiplicative noise model can be rewritten as a sum of the noise free image and a signal dependent additive noise as

$$z = x \cdot v = x + (v - 1)x,$$  \hspace{1cm} (6)

where $1 = [1, \ldots, 1]^T$. In this setting, when the underlying image contains no textures, the first and the second terms in (6) can be viewed as $x_p$ and $x_t$, respectively.

3.2. Optimization

Various methods can be used to obtain the solution of (4). In this paper, we adopt an iterative shrinkage algorithm known as the Separable Surrogate Functionals (SSF) to solve the separation problem [14], [15]. Figure 2 summarizes the different steps of the SSF algorithm, where $S_\lambda(x) = \text{sign}(x)(|x| - \lambda)_+$ is the element-wise soft-thresholding operator with threshold $\lambda$, $(\cdot)_+$ denotes the function $\max(a, 0)$ and $H$ denotes the undecimated Haar wavelet dictionary. Note that we have replaced the TV correction term by a redundant Haar wavelet-based shrinkage estimate as this seems to give the best results. This adjustment is applied only to the piecewise smooth component to control the ringing artifacts near the edges caused by the oscillations of the atoms in the dictionary $D_p$. The same adjustment was used in [12] and the substitution was partially motivated by observing the connecting between $TV$ and the Haar wavelet given in [16]. See [14], [15] for more details on the derivation of this algorithm.

![Image](image-url)

**Fig. 2:** The SSF iterative shrinkage algorithm to solve (4).

### 4. EXPERIMENTAL RESULTS

In this section, we present results of our method using the prototype agile-beam LADAR architecture described in Section 2. In our experiments, we use a dictionary corresponding to the undecimated wavelet transform (Daubechies wavelet with 6 vanishing moments) for the piecewise smooth part and the local DCT dictionary for the texture part. We decay the threshold value of $\lambda = \lambda^k$ during each iteration and stop the iterations when $\lambda^k = 2.1\epsilon$, where $\epsilon$ is the additive noise level [17].

![Image](image-url)

**Fig. 3:** Measured data. (a) Objects used for experiments. (b) Reconstructed objects at depth 3 meters. (c) Background clutter.
4.1. Robust Recovery

In this section, we compare the reconstructions obtained by our method with that of the original system (simply by inverting the system matrix induced by the far-field illumination patterns) and using a standard sparsity promoting $\ell_1$-minimization algorithm where the identity basis is used as the sparsifying transformation. In particular, the following problem is solved to obtain the cross-range element from (1)

$$\hat{x} = \arg \min_x \lambda \|x\|_1 + \frac{1}{2} \|y - Ax\|_2^2.$$  

(7)

We employed a highly efficient algorithm that is suited for large scale applications known as spectral projected gradient (spgl1) algorithm for solving the above problem [18]. Note that we do not reduce the number of measurements in this experiment (i.e. $M = N$).

Reconstruction results are shown in Figure 5. As can be seen from the first row of this figure, simply by inverting the system matrix, one obtains range slides that have high amount of speckle especially in the DCT and Hadamard modes. The sparsity promoting reconstructions (second row) are able to remove some noise but the best reconstructions are obtained by our method (last row). In this figure, we only show the piecewise smooth component as the texture component contains the speckle noise. This experiment clearly shows the significance of using structured representation for obtaining speckle free reconstructions directly from the LADAR measurements.

4.2. CS Recovery

In the second set of experiments, we use the compressive data collected using a random Bernoulli ($p = 0.5$) matrix in our agile-beam LADAR system. We retained only 50% of the measurements (i.e. $M=N/2$). We compare the reconstructed range slices using our method with the reconstruction obtained by solving (7) and by simply inverting ($\ell_2$ reconstruction) the measurement matrix in Figure 6. As expected, the $\ell_2$ reconstruction completely fails to properly reconstruct the range slice. The $\ell_1$-minimization algorithm does reconstruct the range slice properly, however, it suffers from high amount of speckle. In contrast, our method is able to not only reconstruct the objects properly but it is also able to remove the speckle from the range slice.

5. CONCLUSION

In this paper, we studied the problem of directly reconstructing two separate components of LADAR range slices from the LADAR measurements. The approach is based on an optimization formulation that can be solved in the form of thresholding iteration scheme. One of the important advantages of this approach is that the separate components are more amenable to classification tasks yet most importantly solving for these components individually provides an improved reconstruction fidelity. Part of the improvement comes from being able to estimate better the salient and speckle elements. This is because these components are being estimated using separate dictionary that is best adapted to these features unlike previous formulation of this LADAR image recovery problem. More analysis on the artifacts introduced by the random under-sampling in terms of the point spread function and the sensitivity of our algorithms to different amount of noise and sparsity levels will be discussed elsewhere.
6. REFERENCES


