COMPUTATIONAL AGILE BEAM LADAR IMAGING

Arthita Ghosh[1], *Vishal M. Patel*[2] *and Michael A. Powers*[3]

[1]Center for Automation Research, University of Maryland, College Park, MD 20742
[2]Rutgers, The State University of New Jersey, 723 CoRE, 94 Brett Rd, Piscataway, NJ 08854
[3]U.S. Army Research Laboratory, 2800 Powder Mill Road, Adelphi, MD 20783

arthita@umd.edu, vishal.m.patel@rutgers.edu, michael.a.powers70.civ@mail.mil

ABSTRACT

A LAser Detection And Ranging (LADAR) apparatus obtains range information from a three dimensional scene by emitting laser beams and collecting the reflected rays from target objects in the region of interest. The Agile Beam LADAR concept makes the measurement and interpretation process more efficient by a software-defined architecture that leverages Computational Imaging principles to this end. Using these techniques, we show that, the process of object identification and scene understanding can be accurately performed in the LADAR measurement domain thereby rendering the efforts of pixel based scene reconstruction superfluous.

Index Terms— Agile-beam LADAR imaging, object recognition, imaging.

1. INTRODUCTION

In the realm of three dimensional sensing, LAser Detection And Ranging (LADAR) is a well known technology that provides a robust and efficient way to effectively construct 3D scene by using laser light to measure distance to various objects in the scene [1]. The range of the target is measured using the time between the firing of the laser beam and receiving the reflected rays from the target object. Transverse features of the target are resolved by using some form of beam scanning (e.g. a spot or a line) by the LADAR systems. Three-dimensional image of the region of interest is thereafter reconstructed. Other systems like Flash LADAR use a more uniform beam to illuminate objects and an array of detectors to resolve transverse features. The Agile beam LADAR concept utilizes temporal (i.e. a pulse along the direction of propagation) and spatial (i.e. a certain beam pattern in the plane perpendicular to direction of propagation) illumination modulation information to measure a series of inner products and thereafter reconstruct a three dimensional image of the objects in the scene and/or perform object recognition.

This work was sponsored by the U. S. Army Research Laboratory under the Robotics Collaborative Technology Alliance Program, Cooperative Agreement W911NF-10-2-0016 and by Cooperative Agreement 40212299 from General Dynamics.

The Agile beam LADAR concept is particularly useful in robotic perception of three dimensional world, to minimize size, weight, power and cost simultaneously. To this end we have explored the applications of system design and computational imaging algorithms. Emerging technologies in nanophotonic optical phased array devices [2] suggest that future LADAR systems will closely resemble present day electronically scanned RADAR systems and that there is a considerable gain in transferring operations currently implemented in optical hardware to software systems running on digital electronic computers.

An important feature of the Agile Beam LADAR concept is the incorporation of computational approach both in the measurement apparatus (implemented by analog optical operation of lenses and focal planes) as well as the subsequent interpretation process (performed on digital electronic computers). This software-based architecture combining the measurement and interpretation stages has led to opportunities for more efficient measurement and interpretation techniques than simply in the pixel basis rigidly defined in optical hardware. Instances of such technologies include the Rice single-pixel camera [3] which leverages Compressive Sensing (CS) concepts for measurement efficiency and also supports advantageous and cost effective hardware architectures. Some of the other compressive LADAR systems include [4], [5], [6].

In several robotic perception applications, reconstruction of the three dimensional scene on a pixel basis is more than what is necessary, as long as measurement and identification of objects and surfaces can be effectively performed in an alternate basis (e.g. smashed filter approach [7]). Thus avoiding reconstruction without necessarily degrading the quality of perception can help us save computational time and cut down on resource, size and cost requirements. Moreover, incorporation of prior knowledge about the environment in smart utilization of the perception sub-system in a robotic system can simplify and increase the efficacy of interpretation and help with tuning optical parameters.

In this paper, we have explored the idea of object recognition in alternate domains (rather than one in the pixel basis).

Our findings suggest that it is possible to recognize the objects directly in the measurement domain with sufficient levels of accuracy and even attaining comparable performance with the original domain. Our approach saves computational time and resources by avoiding the operations related to reconstruction.

The rest of the paper is organized as follows. Section 2 presents a description of the Agile Beam LADAR concept and the reconstruction process. Section 3 describes how object recognition can be done in the measurement domain. We discuss our experiments and the results obtained in those investigations. Section 4 concludes the paper with a brief summary of the observations and their implications.

2. AGILE BEAM LADAR IMAGING

The agile beam imaging architecture we used in this work is shown in Figure 1. A linearly-polarized transmitter beam at 1550 nm, shown in blue, is generated by a 2-ns pulsed fiber laser and collimated using a multi-element lens (not shown). A mirror is positioned to align the beam to the center of the spatial light modulators (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 a so-called 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.

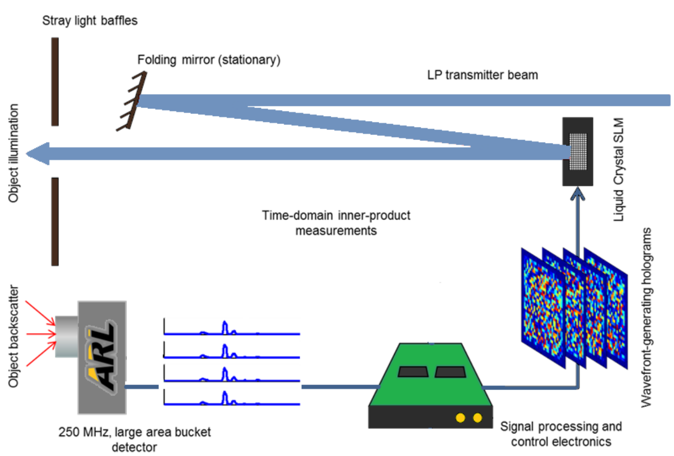
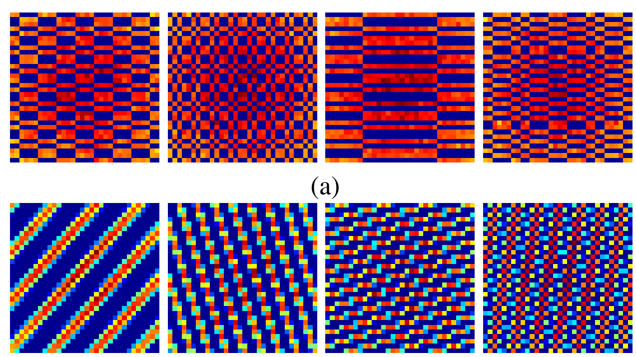


Fig. 1: Block diagram of a prototype agile-beam LADAR architecture.

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-by-n-by-T 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, $32 \times 32 \times 301$ image data cubes were 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 32×32 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 32×32 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 (DCT) matrix and another mode with 1024 digital holograms that project patterns of the rows of the Hadamard transform matrix [8]. Sample patterns are shown in Fig. 2. 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.



(a)

(b)

Fig. 2: Sample patterns used in our experimental prototype. (a) Hadamard (b) DCT.

2.1. Image Recovery

Assuming that we know the range of interest a priori, the agile beam LADAR measurement process can be written as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^N$, $N = n^2$ is the cross-range element, $\mathbf{A} \in N \times N$ is the measurement matrix and $\mathbf{b}$ is the additive measurement noise. Given $\mathbf{y}$ and $\mathbf{A}$, the general problem is to recover $\mathbf{x}$. Direct inversion of the matrix in (1) leads to noisy outputs [9]. As a result robust solutions are usually sought. Assuming that the scene or object of interest is sparse, the following sparsity promoting ℓ_1-minimization algorithm can be used to robustly recover $\mathbf{x}$

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \lambda \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2. \tag{2}$$

We employed a highly efficient algorithm that is suited for large scale applications known as spectral projected gradient (spgl1) for solving the above problem [10].

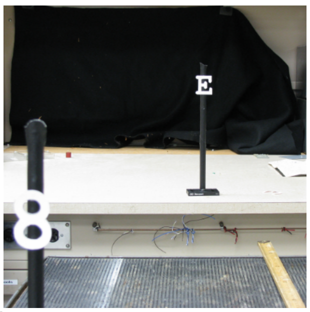


Fig. 3: Objects used for collecting data.

Figure 3 shows two shapes, "8" and "E", that were used for collecting data using the experimental prototype. These shapes were cut from white cardboard and attached to a black post. The first and second shapes were placed about 3 and 4 meters away from the sensor within its field of view, respectively. Peaks in the time histograms were used to identify the slice containing the objects. Peaks at 184 and 229 correspond to the foreground object "8" and background object "E", respectively. The peak at 265 represents the back wall.

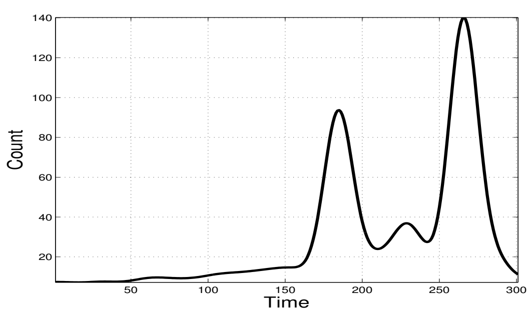
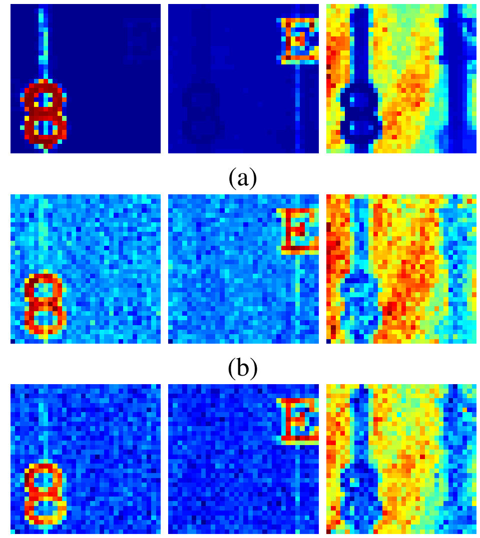


Fig. 4: Timing histogram. Peaks at 184 and 229 represent the foreground object "8" and background object "E", respectively. The peak at 265 represents the back wall.

Reconstructions corresponding to different modes are shown in Figure 5. As can be seen from this figure, the reconstruction quality depends on the mode used for data collection. Furthermore, as the reconstruction procedure involves solving an ℓ_1-minimization problem, it can be very time consuming making it not suitable for real-time processing. For instance, in robotics applications, one might be interested in recognizing objects seen by the LADAR in real time. For this application, reconstructing an image from the LADAR measurements then recognizing the object may not be feasible.



(a)

(b)

(c)

Fig. 5: Measured data. From left to right: reconstructed foreground object "8", reconstructed background object "E" and background clutter. (a) Raster scan. (b) Hadamard. (c) DCT.

3. MEASUREMENT DOMAIN OBJECT RECOGNITION

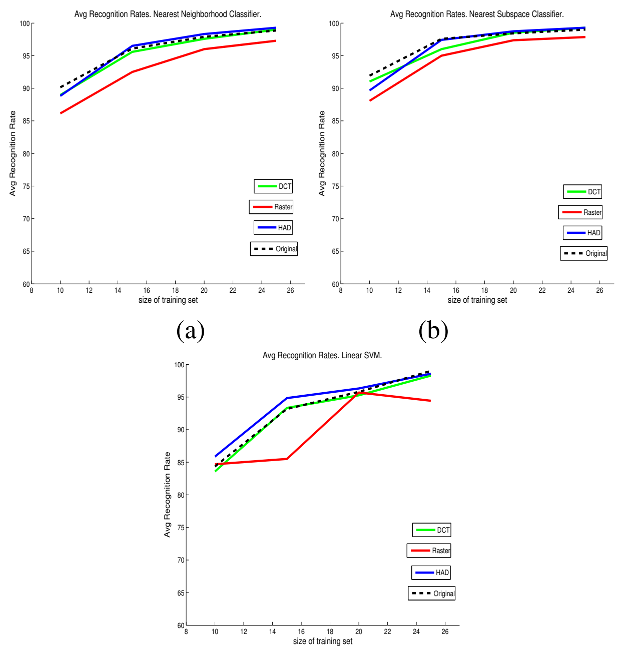
In this section, we investigate whether it is possible to recognize objects directly in the LADAR measurement domain. For the experiments, we use the Binary Alphadigits dataset. [1] The Binary Alphadigits dataset contains binary digits of 0 through 9 and capital A through Z. Each digit is of size 20×16. There are 39 examples of each class. For the experiments in this paper, we used only 10 classes corresponding to digits 0 through 9. We simulated the LADAR measurements using the measured point spread functions (psf) from our agile beam LADAR prototype. In order to simulate the measurements, we first resized the digits to 32×32 and then convolved the digits with the measured psfs. Optimization problem (2) was solved to reconstruct the image from the LADAR measurements.

We compared the performance of three well known classifiers - nearest neighbor (NN), nearest subspace (NS) and linear SVM on this data. We studied the effect of varying the training dataset size on the performance of the classifiers. We took various combinations of training and test datasets and obtained average accuracy rates for various ratios of training set vs. test set size. Moreover we compared the performance achieved in the measurement domains with the figures attained in the original domain of the binary numeric data.

The plots in Figure 6 show the average performance of the classifiers (Linear SVM, NN and NS) in various types of measurement domains. The broken lines in black show the average performance achieved for corresponding training sets

[1] Available at http://www.cs.toronto.edu/~roweis/data.html

in the original binary numeric dataset. While good performance is achieved in most cases, performance in the raster scan mode is uniformly dominated by the performances in the Hadamard and DCT domains. Moreover, the accuracy in these domains are nearly as good, or in some cases slightly better than the original domain. Especially, when the training set size is more than 20, using all three classifiers, we observed similar accuracy levels in the original, Hadamard and DCT measurement domains. These results indicate the effectiveness of classification in the measurement domain thereby bypassing the computational effort in reconstruction of the scene.

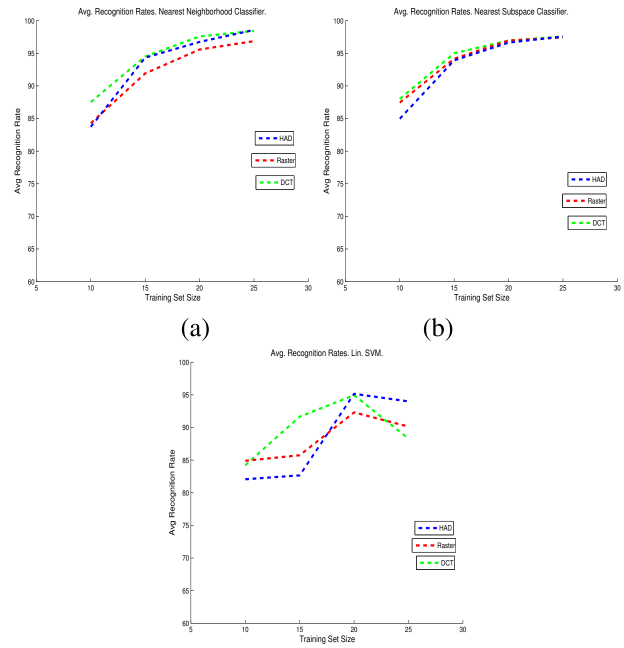


(a) (b)

(c)

Fig. 6: Performance comparison of different classifiers in the measurement domain and the original pixel domain. (a) Nearest neighbor. (b) Nearest subspace. (c) Linear SVM.

In another set of experiments, we investigated the classification performance in the case where we inverted the transformation of the data in the Hadamard and DCT domains to the visual domain. We compared the performance with accuracy level achieved in classifying the data in the raster domain. The plots in Figure 7 indicate that inversion of transformation from the Hadamard and DCT domains result in equally good classification performance than classification in the raster domain in the case of NN and NS classifiers. However these values, again, are similar to the accuracy levels achieved by direct classification in the Hadamard and DCT domain which can be observed by comparing the corresponding (blue and green) curves in Figure 6 and Figure 7 for the particular type of classifier. For example, in the case of our small dataset with 250 training and 110 test samples the average time for NN-classification in the measurement domain was 0.5050 while that for the inverted domain was 0.5247 which showed a 3.9% increase. In the case where we have more samples to invert, the inversion time will be higher. Therefore, the observations made from the above experiments suggest that LADAR data measurement in the DCT or Hadamard domains followed by classification in the same domain itself yields very good performance and cuts down the requirement of computational time and resources.



(a) (b)

(c)

Fig. 7: Performance comparison of different classifiers when the images are recovered by inverting the sensing matrix. (a) Nearest neighbor. (b) Nearest subspace. (c) Linear SVM.

4. CONCLUSION

In this paper, we studied whether is it possible to directly recognize objects in the LADAR domain. Empirical results obtained from experiments with three different classifiers show that indeed it is possible to recognize objects in the measurement domain without explicitly reconstructing them. Classification accuracy in the measurement domain is as good as the original and reconstructed domains. This speeds up the process of interpretation and perception of the three dimensional scene and does away with the necessicity of the reconstruction. In future work, we will investigate the mathematical principles behind the effectiveness of this approach in greater detail.

5. REFERENCES

[1] B. Schwarz, “Lidar: Mapping the world in 3d,” *Nature Photonics*, vol. 4, no. 7, pp. 429–430, July 2010.

[2] J. Sun, E. Timurdogan, A. Yaacobi, E. S. Hosseini, and M. R. Watts, "Large-scale nanophotonic phased array," *Nature Photonics*, vol. 493, no. 7431, pp. 195–199, Jan 2013.

[3] M.F. Duarte, M.A. Davenport, D. Takhar, J.N. Laska, Ting Sun, K.F. Kelly, and R.G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, March 2008.

[4] G. A. Howland, P. B. Dixon, and J. C. Howell, "Photon-counting compressive sensing laser radar for 3d imaging," *Appl. Opt.*, vol. 50, no. 31, pp. 5917–5920, Nov 2011.

[5] A. Colao, A. Kirmani, G. A. Howland, J. Howell, and V. K Goyal, "Compressive depth map acquisition using a single photon-counting detector: Parametric signal processing meets sparsity," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 96–102.

[6] A. Kirmani, A. Colao, F. Wong, and V. K Goyal, "Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor," *Optics Express*, vol. 19, no. 22, pp. 21485–21507, Oct 2011.

[7] Mark A. Davenport, Marco F. Duarte, Michael B. Wakin, Jason N. Laskar, Dharmpal Takhar, Kevin F. Kelly, and Richard G. Baraniukr, "The smashed filter for compressive classification and target recognition," in *Proc. SPIE Computational Imaging*, 2007.

[8] Eugen Romasew, Joachim Barenz, and Hans Dieter Tholl, "Hadamard camera for 3d imaging," in *Proc. SPIE*, 2007, vol. 6737, pp. 67370A–67370A–8.

[9] V. M. Patel and M. A. Powers, "Structured representation-based robust agile-beam LADAR imaging," in *IEEE International Conference on Image Processing*, 2015.

[10] E. van den Berg and M. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2009.