Directionally Controlled Time-of-Flight Ranging for Mobile Sensing Platforms

Zaid Tasneem   Dingkang Wang   Huikai Xie   Sanjeev J. Koppal*

University of Florida

Abstract—**Scanning time-of-flight (TOF) sensors obtain depth measurements by directing modulated light beams across a scene. We demonstrate that control of the directional scanning patterns can enable novel algorithms and applications. Our analysis occurs entirely in the angular domain and consists of two ideas. First, we show how to exploit the angular support of the light beam to improve reconstruction results. Second, we describe how to control the light beam direction in a way that maximizes a well-known information theoretic measure. Using these two ideas, we demonstrate novel applications such as adaptive TOF sensing, LIDAR zoom, LIDAR edge sensing for gradient-based reconstruction and energy efficient LIDAR scanning. Our contributions can apply equally to sensors using mechanical, opto-electronic or MEMS-based approaches to modulate the light beam, and we show results here on a MEMS mirror-based LIDAR system. In short, we describe new adaptive directionally controlled TOF sensing algorithms which can impact mobile sensing platforms such as robots, wearable devices and IoT nodes.**

I. INTRODUCTION

Vision sensors that recover scene geometry have innumerable robotic applications. Recently, a new wave of time-of-flight (TOF) depth sensors have transformed robot perception. These sensors modulate scene illumination and extract depth from time-related features in the reflected radiance, such as phase change or temporal delays. Commercially available TOF sensors such as the Microsoft Kinect [16] and the Velodyne Puck [12], have influenced fields such as autonomous cars, drone surveillance and wearable devices.

Creating TOF sensors for personal drones, VR/AR glasses, IoT nodes and other miniature platforms would require transcending the energy constraints due to limited battery capacity. Recent work has addressed some aspects of TOF energy efficiency with novel illumination encodings. For example, by synchronizing illumination patterns to match sensor exposures [1], low-power reconstruction can occur for scenes with significant ambient light. Additionally, spatio-temporal encodings have been shown to be efficient for both structured light illumination [30] and TOF illumination as well [29].

In this paper, we demonstrate new efficiencies that are possible with *angular* control of a TOF sensor. We demonstrate this with a single LIDAR beam reflected off a micro-electromechanical (MEMS) mirror. The voltages that control the MEMS actuators allow analog (continuous) TOF sensing angles. As a modulator, MEMS mirrors have well-known advantages of high-speed and fast response to control [32].

Our designs provide a new frame work to exploit directional control for depth sensing for applications relevant to small robotic platforms. Our experiments use a pulse-based LIDAR, but the algorithms can easily be extended to continuous wave systems, as well as any other method for modulation such as mechanical [12] or opto-electronic [33].

Our contributions are:

- We provide imaging strategies for directional control of TOF samples, with a particular focus on the angular support of the sensing beam. We demonstrate, through real experiments and simulations, that deblurring the measurements using the sensor's angular support can recover high-frequency edges, correct non-uniform sampling and is robust through wide field-of-view (FOV) distortions.
- We discuss a information-theoretic-based control algorithm for the MEMS mirror, in order to decide which scan to generate, given the previous measurements. By changing the cost function in the control algorithm, we can create energy-efficient TOF 3D sensing, where the algorithm places samples where they are most needed. Our method optimizes 3D sensing accuracy along with physical constraints such as range-derived power consumption, motion of objects and free space coverage.
- We demonstrate all of our algorithms on a real sensor, and show additional applications that are relevant for small robotic platforms, such as LIDAR zoom, which allows the controller to investigate interesting regions in the scene, as well as gradient-based estimation, which allows a constrained system to place its samples along edges, and reconstructs the scene post-capture.

II. RELATED WORK

TOF imaging and adaptive optics: Efficient TOF reconstruction is possible in the face of global illumination by encoding phase frequencies [11] or through efficient probing [29]. Synchronization with camera exposure has allowed for reconstruction in the face of strong ambient light [1]. Transient imaging is possible using ultra-fast lasers [42], and has recently been demonstrated using mobile off-the-shelf devices [14]. We focus on 3D reconstruction and show that directional control can allow for novel types of efficiencies in sampling and energy consumption. Finally TOF sensors for long range sensing through atmosphere uses fast adaptive optics to remove atmospheric turbulence effects [3, 41], whereas we target scene-adaptive sensing for autonomous systems.

(I) Ray diagram of our sensor design

(II) Experimental setup
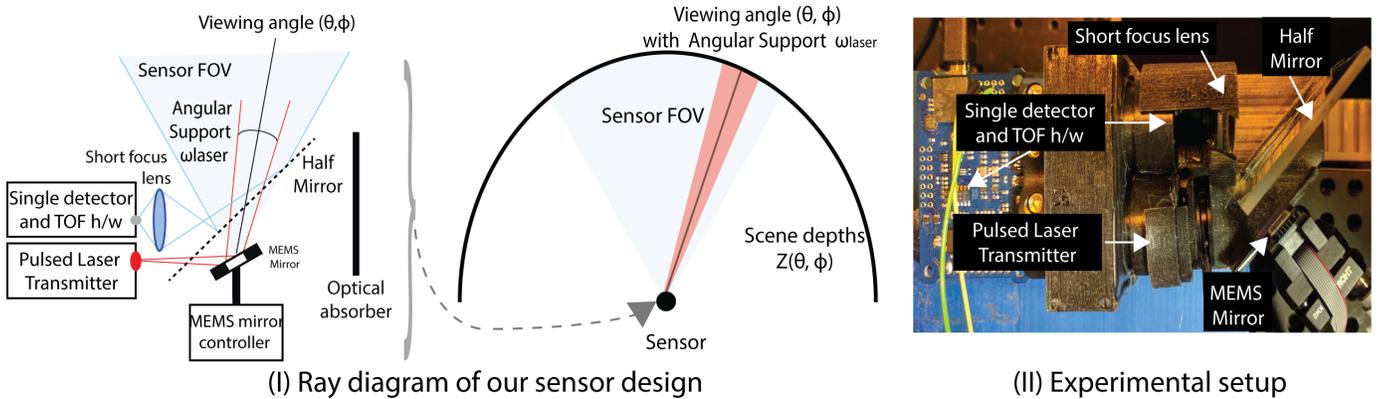
Fig. 1. Ray diagram: In (I) we show the co-located ray-diagram of the pulsed LIDAR that is modulated by a MEMS device. In (II) we show our actual setup, with a Lightware LIDAR and a Mirrorcle MEMS mirror.

Adaptive sampling in 3D models: Adaptive sampling techniques from signal processing [34] are used extensively for efficient mesh representations of computer generated scenes [39, 4]. In robotics and vision, information theoretic approaches are used to model adaptive 3D sensing for SLAM and other applications [40, 5, 15]. In this paper, we are interested in adaptive algorithms for LIDAR sensors that take into account physical constraints such as the power expanded on far away objects or on objects moving out of the field-of-view. We demonstrate the balancing of such efficiency goals with 3D reconstruction quality.

MEMS mirrors for vision: The speed and control of MEMS mirrors have been exploited for creating imperceptible structured light for futuristic office applications [35] and interactive-rate glasses-free 3D displays [17]. MEMS mirror modulated imaging was introduced through reverse engineering a DLP projector [26] for tasks such as edge detection and object recognition. Coupling a DLP projector with a high-speed camera allows for fast structured light and photometric stereo [20]. Adding a spatial light modulator in front of the camera allows for dual masks enabling a variety of applications [30], such as vision in ambient light. In contrast to these methods, we propose to use angular control to enable new types of applications for 3D imaging. We are able to play off angular, spatial and temporal sampling to allow, for example, increased sampling in regions of interest.

Scanning LIDARs: Most commercially available LIDARs scan a fixed FOV with mechanical motors, with no directional control. MEMS modulated LIDARs have been used by NASA Goddard's GRSSLi [8], ARL's Spectroscan system [38] and Innoluce Inc. [22]. In all these cases, the MEMS mirrors are run at resonance, while we control the MEMS mirror to demonstrate novel imaging strategies. MEMS mirror control is achieved by [19] at Mirrorcle Inc., who track specially placed highly reflective fiducials in the scene, for both fast 3D tracking and VR applications [25, 24]. We do not use special reflective fiducials and utilize sensing algorithms for MEMS mirror control. Finally, in [37] a MEMS mirror-modulated 3D sensor was created, with the potential for foveal sensing, but without the type of adaptive algorithms that we discuss.

III. MEMS-MODULATED LIDAR IMAGING

A MEMS-modulated LIDAR imager has the following advantages:

- The MEMS mirror's angular motion is continuous over its field-of-view (FOV).
- The MEMS mirror can move selectively over angular regions-of-interest (ROIs).

In this section, we discuss some preliminaries that are needed to actualize these advantages in imaging algorithms. We first show how to use the advantage of continuous motion to remove deblurring artifacts. We then discuss how to use the advantage of selective motion to enable TOF measurements that maximize an information theoretic metric.

A. Sensor design and calibration

A MEMS-modulated LIDAR imager consists of a time-of-flight engine and a MEMS modulator, as in Fig 1(I). The engine contains a modulated laser transmitter, a receiving photodetector that measures the return pulses and additional electronics to calculate the time between transmitted and received pulses.

To avoid errors due to triangulation, we co-locate the centers of projection of the transmitter and receiver, as shown in Fig 1(I). Unlike previous efforts, such as [8, 38] we do not co-locate the fields-of-view of the transmitter and receiver - i.e. the MEMS mirror is not our transmitter's optical aperture. This allows us to avoid expensive and heavy GRIN lenses to focus the laser onto the MEMS device. Instead we use a simple, cheap, light-weight short focus thin-lens to defocus the receiver over the sensor's FOV. This introduces a directionally varying map between the time location of the returned pulses's peak, and the actual depth. We correct for this with a one-time calibration, obtained by recovering 36 measurement profiles across five fronto-parallel calibration planes placed at known locations, as shown in in Fig 2.

Current configuration specs. In Fig 1(II) we show our current configuration, where we use an open-source hardware 1.35W Lightware SF02/F LIDAR and a Mirrorcle 3.6mm Al-coated electrostatic MEMS mirror. The LIDAR operates in NIR (905nm) and the MEMS response is broad-band up to
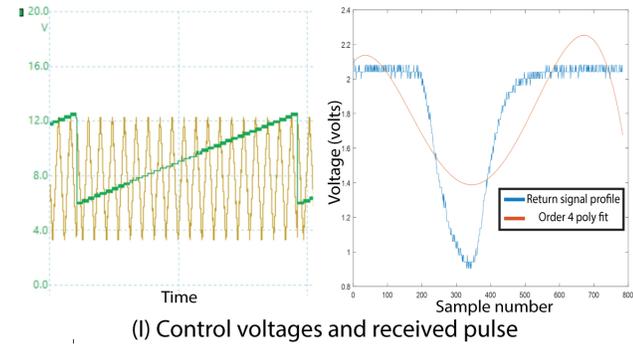
(I) Control voltages and received pulse
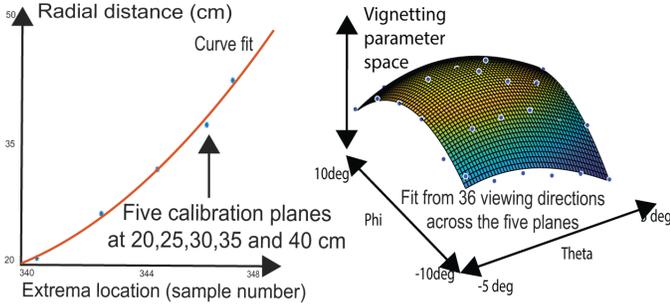
(II) Pulse extrema and vignetting across five calibration planes
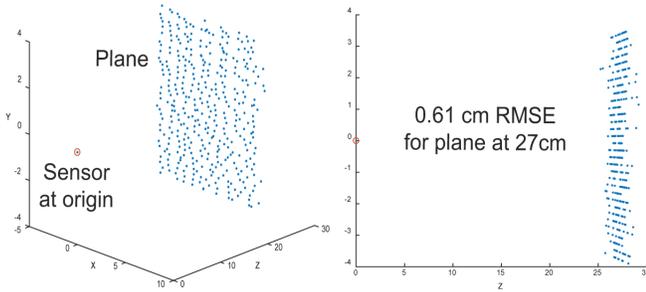
(III) Validation by reconstructing a plane at 27cm

Fig. 2. Calibration: In (I) we show a screenshot of the voltages that control the MEMS mirror pose, as well as an example of the return pulse. We fit a fourth-order polynomial to the return pulse to detect the extrema location. In (II) we show how to map this extrema location to depth in cm, by collecting data across five planes at known depths. This calibration also involves a vignetting step, to remove effects in the the receiver optics. In (III) we validate our sensor by reconstructing a fronto-parallel plane at 27cm, showing a standard deviation of 0.61cm (i.e. almost all points are measured in a ± 1.5cm error range). In the current configuration, the FOV is $\approx 15°$ and the range is 0.5m.

long wave IR (14 μm). As shown in Fig 2(I), voltages from an oscilloscope control the MEMS mirror direction, and the synchronized received pulses are inverted.

The short focus lens introduces an interesting trade-off between FOV, range and accuracy. At the extreme case, with no lens, our FOV reduces to a single receiving direction with the full range of the LIDAR (nearly 50m). As we increase the FOV, and the defocus, the SNR received at the transducer decreases, reducing range. While we can compensate with increased gain, this introduces noise and reduces accuracy. In this paper, we traded-off range for accuracy and FOV, and our device has a FOV of $\approx 15°$, a range is 0.5m and is set at the lowest gain (highest SNR).

The Lightware LIDAR sampling rate is 32HZ, which, in

this paper, restricts us to static scenes. We prefer this LIDAR, despite the low rate, since it allows for raw data capture, enabling design-specific calibration. We perform multiple scans of the static scene, averaging our measurements to improve accuracy. After the calibration in Fig 2(II) we reconstruct a plane at 27cm (not in our calibration set) and obtained standard deviation is 0.61cm (i.e. almost all points are measured in a ± 1.5cm error range), as shown in Fig 2(III).

The total weight of our system is approximately 500g; however most of that weight ($350g$) is contained in a general purpose oscilloscope and MEMS controller, and it would be trivial to replace these with simple, dedicated circuits (67g Lidar, 187g oscilloscope, 74g enclosure, 10g optics and 147g MEMS controller).

B. Directional control of TOF sensing

Voltages over the MEMS device's range physically shift the mirror position to a desired angle, allowing for range sensing over the direction corresponding to this angle. Let the function controlling the azimuth be $\phi(V(t))$ and the function controlling elevation be $\theta(V(t))$, where V is the input voltage that varies with time t. W.l.o.g, we assume a pulse-based system, and let the firing rate of the LIDAR/TOF engine be $\frac{1}{T_f}HZ$, or T_f seconds between each pulse. Therefore, the nth measurement of the sensor happens along the ray direction given by the angles $(\theta(V(n\ T_f)), \phi(V(n\ T_f)))$.

The world around a miniature vision sensor can be modeled as a hemisphere of directions (Fig 1(I) center), i.e. the plenoptic function around the sensor is an environment map parameterized by the azimuth and elevation angles. Just as conventional imagers are characterized by their point-spread function (PSF [9]), miniature vision systems are characterized by their angular support ω [21]. For miniature active scanning TOF systems, the angular spread of the laser beam determines the angular support, which we term as ω_{laser} in Fig 1.

C. Deblurring TOF measurements over the FOV

Each sensor measurement occurs across the laser's dot size, given by the angular support ω_{laser}. Let us now define the separation *between* measurements in angular terms, as ω_{diff}. For many commercial LIDARs, such as the Velodyne HDL-32E, the measurement directions are further apart than the angular support; i.e. $\omega_{diff} \geq \omega_{laser}$.

For our system, the measurement separation ω_{diff} depends on the differential azimuth-elevation, given by $\omega_{diff} = \delta\phi\ \delta\theta\ sin(\phi)$, where ϕ and θ were defined previously.

MEMS-modulation allows almost any angle inside the sensor's FOV. Therefore, if the measurements satisfy the inequality $\omega_{diff} \leq \omega_{laser}$, then the measurements are correlated.

Therefore, we have defined the inequality that transforms our sensor into one identical to fixed resolution imagers. If this inequality is satisfied, then the rich body of work in vision using controllable PSFs can be applied here, including image deblurring [36], refocusing [28], depth sensing [23] and compressive sensing [7].
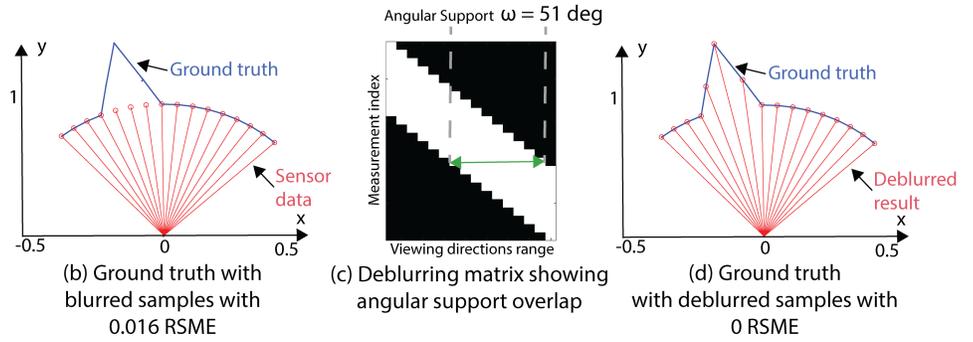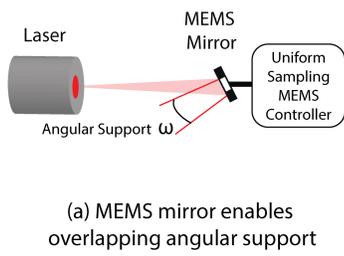
(I) Deblurring 1D simulated data with uniform sampling

(a) MEMS mirror enables overlapping angular support

(b) Ground truth with blurred samples with 0.016 RSME

(c) Deblurring matrix showing angular support overlap

(d) Ground truth with deblurred samples with 0 RSME
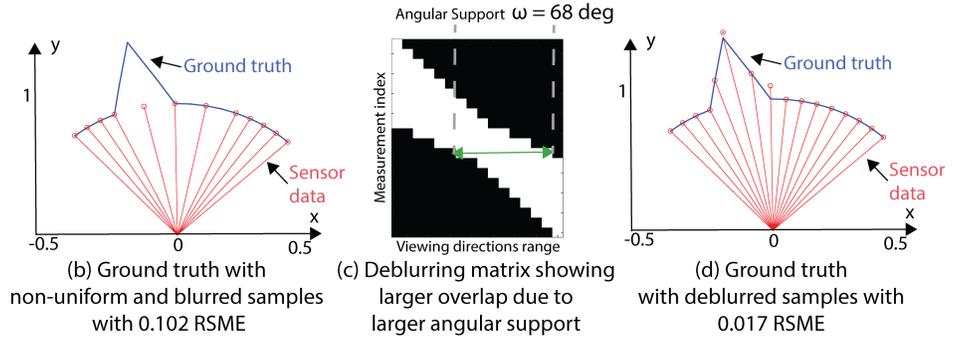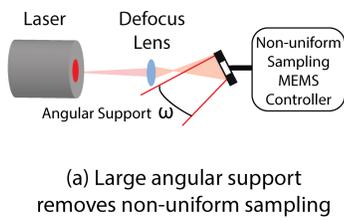
(II) Deblurring 1D simulated data with non-uniform sampling

(a) Large angular support removes non-uniform sampling

(b) Ground truth with non-uniform and blurred samples with 0.102 RSME

(c) Deblurring matrix showing larger overlap due to larger angular support

(d) Ground truth with deblurred samples with 0.017 RSME
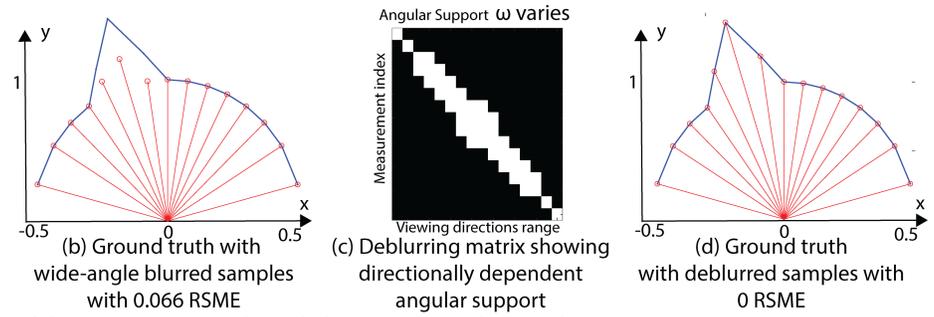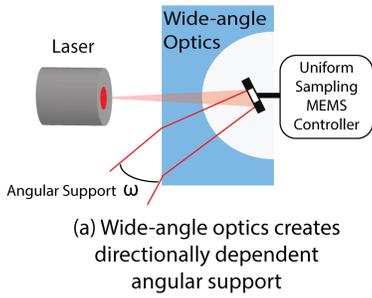
(III) Deblurring 1D simulated data with wide-angle optics

(a) Wide-angle optics creates directionally dependent angular support

(b) Ground truth with wide-angle blurred samples with 0.066 RSME

(c) Deblurring matrix showing directionally dependent angular support

(d) Ground truth with deblurred samples with 0 RSME
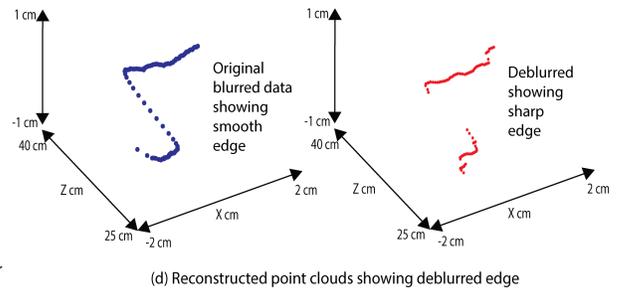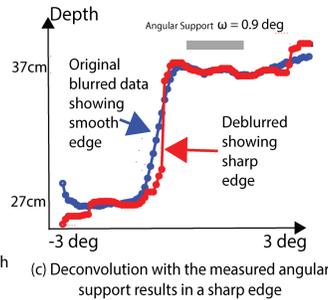
(IV) Deblurring real 3D edge data

(a) Real experiment showing scanning of a depth edge

(b) NIR image of laser dot with angular support of 0.9 deg

(c) Deconvolution with the measured angular support results in a sharp edge

(d) Reconstructed point clouds showing deblurred edge

Fig. 3. Deblurring using angular support: In (I-III) we show simulations of pulse-based LIDAR with 2% noise on a simple 2D circular scene with a sharp protrusion. In (Ia-b) the native laser dot size blurs the scene equiangularly, resulting in a banded matrix I(c) which is invertible I(d). In II (a-b) a larger angular support is created with additional optics, for a sensor with non-uniform angular sampling. The matrix is still invertible, and can be used to resample the scene uniformly II(c-d). (III) shows the effect of wide-angle optics, modeled from the refractive optics in [21, 43], where deblurring is still successful. Finally (IV) shows real experiments across a depth edge. Without deblurring, depth measurements are hallucinated across the gap.

a. Base scene scan | b. Geometric complexity visualization | c. One of many ROI clustering choices | d. Candidate scans in descending CSQMI order | e. Scans with the worst and best score
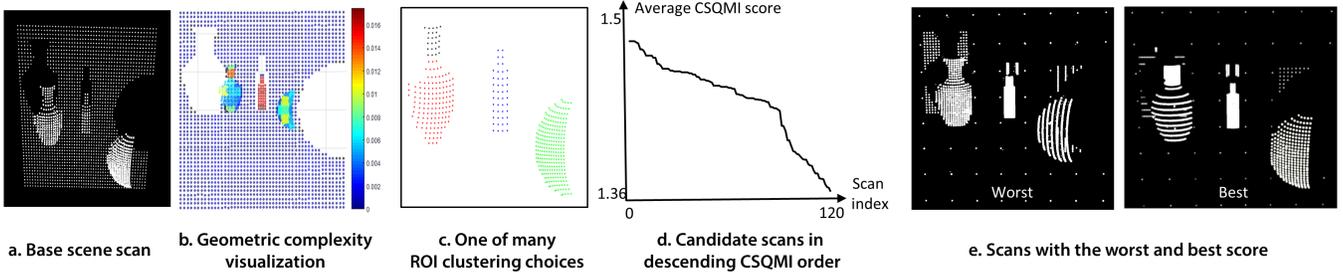
Fig. 4. Simulation of adaptive LIDAR: In (a) we show a base scan of a scene with three simple objects. This base scan is tesselated into boxes which are scored according to some desired metric. In (b) we show a geometric score based on the residual of 2D principal component analysis of points in a box. Note the background plane has a low score. This score is used with multiple values of cluster number k to generate many region-of-interest (ROI) segmentation, as in (c). The ROIs globally direct the MEMS mirror to scan longer in regions with a higher average box score. For each ROI segmentation, many candidate scans are generated by varying scan parameters such as phase, shape, directionality, etc, whose CSQMI scores are shown in (d). We show the highest and lowest average CSQMI scores of these scans, and the highest scan's MEMS mirror motions would be the actual trajectories scanned next.

As an example, in Fig 3, we show noisy simulations of a 2D toy scene where a MEMS modulated LIDAR is shown to be scanning a circular scene with a sharp discontinuity. The angular support ω_{laser} is shown to be much larger than the differential measurement angle ω_{diff}, and therefore the direct measurements blur the high frequency information in the sharp discontinuity in Fig 3 I(b).

Assuming the intensity across the laser dot is uniform, we can represent the angular support for any particular MEMS mirror position as an indicator vector along viewing direction. We concatenate these binary indicator vectors across MEMS mirror directions over the entire FOV to give a measurement matrix $\mathbf{B}$ shown in Fig 3 I(c). In Fig 3 I(c), the rows of the matrix $\mathbf{B}$ are the indices of different measurements, and the columns cover discrete viewing directions across the FOV. Any measurement collects information across the angular support in this FOV, given as white, and ignores the rest, shown as black.

Therefore the measured, received pulses at the sensor are given $y = \mathbf{B}z$, where $\mathbf{z}$ is a vector of the ground-truth received pulses for an ideal angular support ω_{laser}. Recovering the $\mathbf{z}$ is a deblurring problem and we apply non-negative least squares to obtain measurements as shown in Fig 3 I(d), with zero RMSE error.

We note that the angular support of the sensor is constant across viewing direction, because it is simply the angular spread of the laser being reflected off the MEMS mirror. This results in a near-perfect banded diagonal matrix in Fig 3 I(c), which is invertible. The angular spread can be affected by adding laser optics, as shown in Fig 3 (II)a, where the angular support ω_{laser} is increased.

This would be necessary if the maximum angular spread between measurements is much larger than the original angular support ω_{laser}, due to non-uniform MEMS mirror control. In fact, such control occurs naturally with MEMS devices driven by linear signals, since the MEMS device's forces follow Hooke's law of springs [44]. In Fig 3 II(c) the non-uniform and blurred measurements result in a banded matrix with varying band-width. The ground-truth recovered is both accurate and with the desired uniform density sampling.

Finally, consider the effect of a wide-angle optical system,

such as a fish-eye lens. This would shear the diagonal band in the matrix, where extreme angles would integrate large portions of the field of view, which samples closer to the optical axis would show finer angular resolution. The smooth motion of the MEMS mirror allows us to invert or redistribute the samples across the field-of-view, removing wide-angle distortion. In Fig 3 III(c) we shown an example of such optics [21, 43] which has been used recently in wide-angle MEMS modulation. Using the equation from [21], we generate the viewing-dependent angular support that creates a blurred version of the scene in Fig 3 III(b) and a non-banded diagonal matrix in Fig 3 III(c). This matrix is geometrically constructed to be invertible since all its values as positive and its trace is non-zero, and allows for recovery of the scene in Fig 3 III(d).

In Fig. 3 (IV) we show a real deblurring result. The scene is two planes at 27cm and 30cm, where the sensing angles follow a great arc in the hemisphere of directions, as shown by the line segment in Fig. 3 (IV)a. We measure the angular support ω_{laser} as $0.9°$ as shown in Fig. 3 (IV)b. Without deblurring, the original measurements result in a smoothed edge, as shown in Fig. 3 (IV)c-d in blue. We use the damped Richardson-Lucy blind deconvolution optimization algorithm that takes our measured angular support as a starting point, as shown in Fig. 3 (IV)c. This results in a strong edge recovery, with fewer incorrect measurements, as shown in red in Fig. 3 (IV)c-d.

D. Adaptive TOF sensing in selected ROIs

The MEMS mirror can modulate the LIDAR beam through a range of smooth trajectories. A unique characteristic of our setup is that we can adapt this motion to the current set of scene measurements. We control the MEMS mirror over the FOV by exploiting strategies used for LIDAR sensing in robotics [18, 5, 40]. In particular, we first generate a series of candidate trajectories that conform to any desired global physical constraints on the sensor. We then select from these candidates by maximizing a local information theoretic measure that has had success in active vision for robotics [5].

Candidate trajectories from global physical constraints To generate a series of candidate trajectories, we encode the scene into regions where the sensing beam should spend more time collecting many measurements, and regions where the
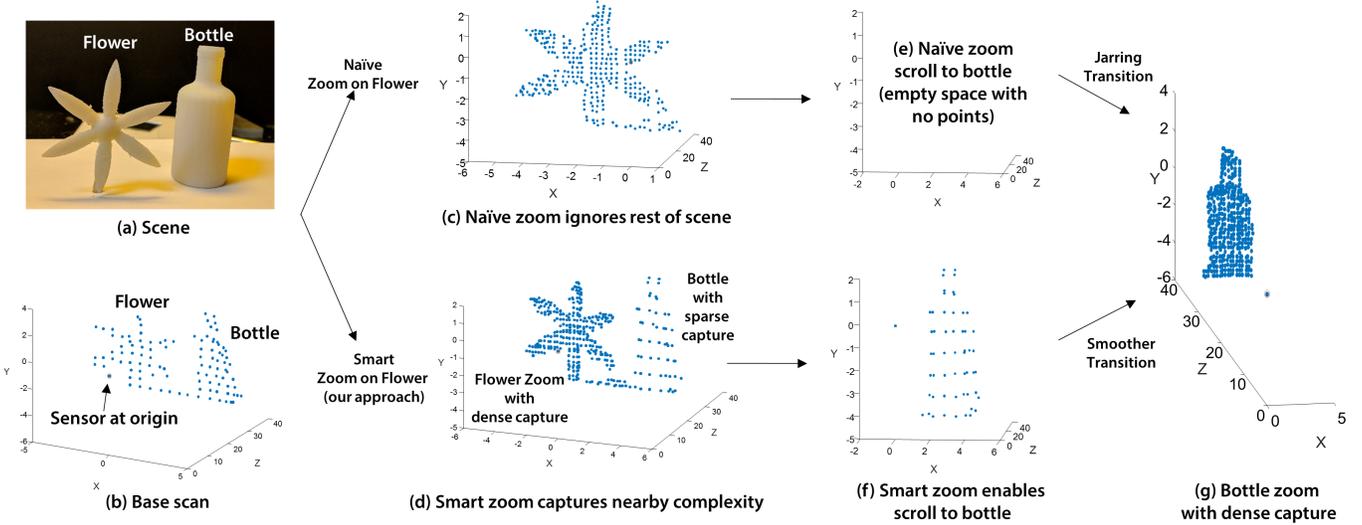
Fig. 5. Our smart LIDAR zoom vs. naive LIDAR zoom: By moving the MEMS mirror to certain regions of interest, we can "zoom" or capture more angular resolution in that desired region. In (a) we show a scene with two objects, and in (b) we show the output of our sensor with equiangular sampling. If the zoom shifts to the flower, then the naive zoom concentrates the samples in the base scan on the flower exclusively, in (c). On the other hand, our smart zoom (d) takes measurements outside the zoom region, depending on neighboring object's complexity and proximity. A naive approach does not visually support scrolling, since other areas of the scene are blank (e). Our smart zoom allows for scrolling to nearby objects that have some measurements (f). This allows a for smoother transition when the zoom shifts to that object (g), when compared to the naive zoom.

beam should move quickly, collecting fewer measurements. We achieve this by clustering the scene into regions of interest (ROI) based on a desired physical metric. In the applications section, we show that different metrics can enable, for example, scanning the scene under the constraint of limited power. Similar metrics can be specified for time or scene complexity.

We first tessellate the current scene scan in three dimensions into bounding boxes $B_i(X_c, Y_c, Z_c, H)$, which contain all points (X, Y, Z) in the current scan such that these lie in a box centered at (X_c, Y_c, Z_c) with side length given by H. We require that a metric M be designed such that $M(B_i) \in \mathbf{R}$. We then apply an unsupervised clustering mechanism, such as k-means, to the set of boxes, where the feature to be clustered from each box B_i is $(M(B_i), X_c, Y_c, Z_c)$. Automatically finding the number of clusters is an open problem in pattern recognition, and while a variety of methods exist to find an optimal k, for simplicity we generate candidate trajectories over a range of cluster centers, from 2 till k_{max}, which we leave as a design parameter.

Each cluster of boxes defines a region of interest (ROI) in the scene. Let us describe the solid angle subtended by the ROI onto the MEMS mirror, indexed by j as ω_j, and let the average metric of all the boxes in the jth ROI be m_j. If there are n samples across the FOV, then our goal is to create a series of voltages $V(t)$, such that the angles generated maximize the following cost function,

$$\max_{V(t)} \Sigma_i^{(n \ T_f)} \Sigma_j^k F(\theta(V(n \ T_f)), \phi(V(n \ T_f)), \omega_j, m_j), \quad (1)$$

where k is the number of ROI clusters in that scan, $\frac{1}{T_f}$ is the firing of the LIDAR/TOF engine and where F is a function that outputs $e^{\alpha m_k}$ if $(\theta(V(n \ T_f)), \phi(V(n \ T_f))$ lie inside ω_k. α

is a user defined weight that controls the impact of the relative score of the different ROIs on the time spent in each ROI. If $m_{total} = \Sigma_j^k m_j$, we pick a weight α such that the time spent in each ROI is proportional to $\frac{m_j}{m_{total}}$.

Note that the above equation does not contain a derivative term $V'(t)$ to enforce smoothness, since we generate only candidate trajectories that conform to physically realizable MEMS mirror trajectories, such as sinusoids, triangular wave functions and raster scans. We generate P such scans $V_p(t)$ and pick the scan that maximizes the sum of the information gain from each scanned ray

$$\max_{V_p(t), m} I(m|x_t) \quad (2)$$

where x_t is the current scan and m is the probabilistic occupancy map of the scene calculated by tessellating the scene into voxels $B_i'(X_c', Y_c', Z_c', H')$, and where the probability of occupancy is given by $e^{-0.5r}$, where r is the radial distance between the voxel center and the nearest measured scan point.

We use a form for $I(m|x_t)$ in Eq 2 derived from the Cauchy-Schwarz quadratic mutual information (CSQMI) for a single laser beam [5]. The expression for CSQMI is reproduced here from [5],

$$\log \Sigma_{l=0}^C w_l \mathcal{N}(0, 2\sigma^2)$$
$$+ \log \Pi_{i=1}^C (o_i^2 + (1 - o_i)^2) \Sigma_{j=0}^C \Sigma_{l=0}^C p(e_j) p(e_l) \mathcal{N}(\mu_l - \mu_j, 2\sigma^2)$$
$$- 2 \log \Sigma_{j=0}^C \Sigma_{l=0}^C p(e_j) w_l \mathcal{N}(\mu_l - \mu_j, 2\sigma^2)$$

$$(3)$$

where C refers to the number of "cells" — voxels intersected by the current laser ray direction, $\mathcal{N}$ and σ define the mean and variance of a Gaussian model of the return pulse, o_i is the probability that the ith cell is occupied, $p(e_j)$ is the

probability that the jth cell is the first occupied cells (with all before being unoccupied) and w_l is a weight defined by $w_l = p^2(e_l)\Pi_{j=l+1}^{C}(o_j^2+(1-o_j)^2)$. Since we have multiple ray directions in each candidate scan, we aggregate each of these to produce a single, overall CSQMI value for that candidate scan and pick the scan with the maximum score.

Simulation example In Fig 4 we show a scene created with BlenSor [10] with three objects in front of a fronto-parallel plane. We start with a equi-angular base scan shown in Fig 4(a) since all directions have uniform prior. We tesselate the scene into boxes B of size $25cm \times 25cm \times 25cm$ and use the residuals of a 2D principal component analysis fit to score the complexity of each box, as in Fig 4(b). Clustering the boxes Fig 4(c) creates regions of interest. Varying the number of clusters and varying scan parameters creates a variety of candidate, each of which have a CSQMI score Fig 4(d). We pick the best such score, as shown in Fig 4(e), where it is contrasted with the worst such scan. Note that, in the best scan, the neck of the vase is captured in detail and the sphere is captured equally densely across θ and ϕ angles.

IV. APPLICATIONS

The imaging framework we have just described allows for directional control of TOF measurements. Using these ideas and our MEMS mirror-based LIDAR sensor, we demonstrate the following novel applications.

A. Smart LIDAR zoom

Optical zoom with a conventional fixed array of detectors involves changing the field-of-view so that the measurements are closer together in the angular domain. The key component of zoom is that new measurements are made, when compared to the initial image.

Intelligent zoom exists for conventional cameras using pan-zoom-tilt transformations [6] and light-fields [2]. Here we demonstrate, for the first time, intelligent LIDAR zoom.

Suppose we are provided with angular supports of n interesting regions of interest in the scene $(\omega_{zoom}^1, \omega_{zoom}^2, ...\omega_{zoom}^n)$ and a corresponding series of importance weights $(w_1, w_2, ..., w_n)$. These could come from another algorithm, say face detection, or from a user giving high-level commands to the system.

We can use these regions and weights to modify the default effect of the adaptive LIDAR framework described in Sect III-D. For example, if a box is contained in ω_{zoom}^i, then we can increase the geometric score in the boxes by a factor determined by the corresponding importance weight w_i. This would increase the amount of time that the sensor spends in the angular support corresponding to the box.

Smart LIDAR zoom has a clear advantage over naive zoom, which would place all LIDAR samples exclusively in ROIs. This because any zoom interface must also offer scrolling. As is known in computer graphics [27], efficient scrolling requires caching motion and data near user viewpoints, to allow for fast rendering for real-time interaction.

a. Energy efficient LIDAR simulation

b. Energy efficient LIDAR with real scenes
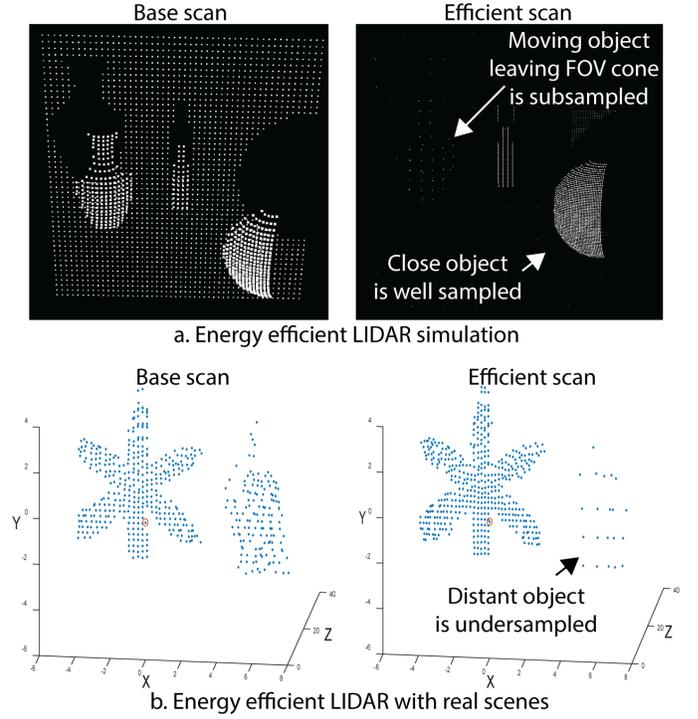
Fig. 6. Energy aware adaptive sampling: We augment our directional control algorithm for adaptive TOF with physical constraints that capture the energy budget of the system. Here we use two constraints; the inverse fall-off of light beam intensity and the motion of scene objects w.r.t to the sensor FOV. In (a) we show simulations of three objects, one of which is give a motion perpendicular to the optical axis (i.e. leaving the sensor FOV cone). Compared to the base scan (left), the effiency scan reduces sampling on objects that move beyond the FOV cone and distant objects, despite their complexity. In (b) we show a real example using our MEMS mirror-based sensor, where, again, distant objects are subsampled.

In Fig 5(a) we show a scene with a 3D printed flower and a bottle. We show a base scan of the scene in Fig 5(b) with equiangular samples. The user places a zoom region of interest around the flower. We show that naive zoom directs the measurements entirely on the flower, with almost zero measurements around the zoomed-in area.

While we have not implemented real-time scrolling, in Fig 5(c-g) we simulate the effect of scroll in naive zoom, showing a jarring transition period in the image, since the measurements suddenly appear in a previously blank image. Instead, our method spreads the samples across the two complex objects in the scene, allowing for a more meaningful transition when scrolling is simulated in Fig 5(g) to the dense scan when the bottle is zoomed. Note, that while the scroll motion is simulated, all the zoom measurements are real measurements from our sensor performing directionally varying sampling, based on the desired zoom area.

B. Energy-aware adaptive sampling

In its current form, the adaptive TOF sensing algorithm only uses a geometric goodness metric. To augment the algorithm for mobile-based platforms, we wish to include multiple, say n, physical constraints into the metric. Therefore we redefine the metric as $M(B_i) \in \mathbf{R^n}$, where B_i is the ith box in the tessellated current scan.

a. Reconstruction of two fronto-parallel planes

b. Top view

c. Capturing high-quality data by averaging samples in edge regions

d. Gradient-based full mesh reconstruction

e. Top view of reconstructed meshes
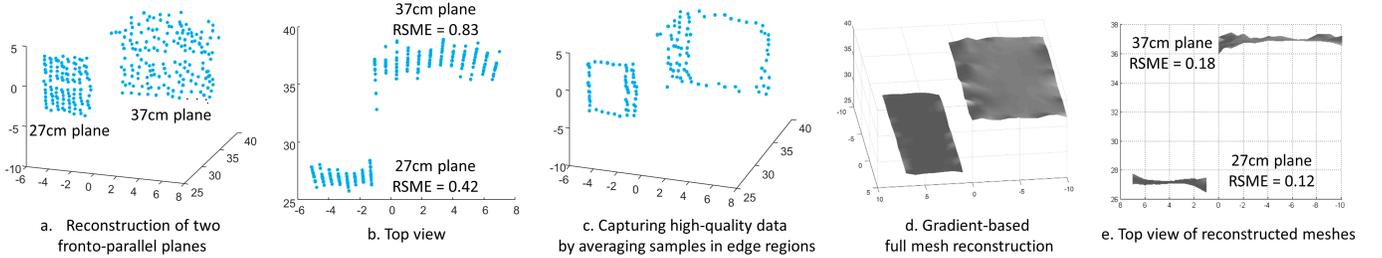
Fig. 7. Gradient-based reconstruction: Directional control allows a capture of the scene, where samples are only made in high-frequency (i.e. edge) regions of the scene. In (a) we see the original scan of the two planes, and (b) illustrates their noise levels. In (c), we directly capture only edge regions, placing the same number of samples as in (a) in high-frequency areas, improving averaging and reducing error. We use these edges with a gradient-based reconstruction algorithm to recover the meshes in (d). Note that the noise levels are significantly reduced, as shown in (e).

To illustrate the redefined metric, we point out differences between adaptive sensing when compared to adaptive sampling literature in image processing and graphics. First, for a given pulse signal and desired SNR, distant objects require more pulses. Therefore, geometric complexity must trade-off with range, and a distant, intricate object may not be sampled at the required resolution, to save energy. Second, temporal relevance matters, and a nearby, intricate object that is rapidly moving out of the field-of-view need not be sampled at high resolution. Third, unlike virtual scenes, free space must be sampled periodically, since new obstacles may emerge. Finally, the sensor's measurement rate, implies finite samples which must be shared across all objects, complex or simple.

The issues of free space and infinite samples are already handled by the adaptive algorithm described in Sect III-D, and we augment it with two new metrics in $M(B_i)$. The first takes into account the dissipation of the laser, and scores distant objects by two-way inverse square reduction in radiance, or $\frac{1}{Z^4}$. The second is simply a scaled version of the object's velocity $\lambda \vec{v}$, where λ is 1 if the direction $\vec{v}$ is contained in the FOV cone of the sensor, and zero otherwise.

Fig. 6 shows both simulated and real scenes, where objects are at different distances from the sensor. In the simulated scene, the vase is given a motion away from the sensor's visual cone. In the first column we see the base scan of the scene, where samples are taken equiangularly. Applying the physical restrictions discussed above and using the adaptive algorithm described in Sect III-D produces the results in the second column, where samples are reduced to save energy consumption and time.

C. Edge sensing for gradient-based reconstruction

Gradient-based methods [31] have had significant impact in vision, graphics and imaging. Given a base scan of the scene, we can focus our sensor to place samples only on regions of high-frequency changes in depth. Placing all our samples in these regions, over the same time it took to previously scan the entire scene, produces more robust data since averaging can reduce noise in these edge regions.

Our goal is to estimate scene depths Z, from a small number of captured depths $\hat{Z}$. A popular solution is to minimize some norm between the numerically computed real and estimated x
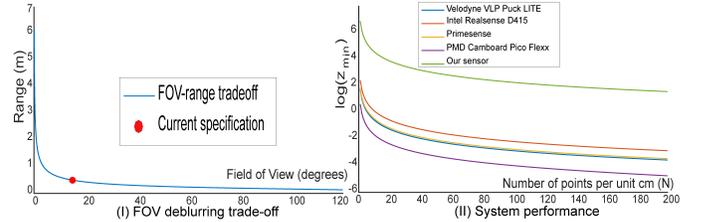
Fig. 8. In (I) we depict the deblurring trade-off in our setup, where increasing FOV results in reduced SNR and range. In (II) we compare our LIDAR zoom to other available depth sensors which have fixed acuity.

and y gradients. Formulating this for our scenario,

$$\min_Z \| \frac{d\hat{Z}}{dx} - \frac{dZ}{dx} \|^2 + \| \frac{d\hat{Z}}{dy} - \frac{dZ}{dy} \|^2. \tag{4}$$

Note that the minimization estimates scene depth Z, which has values outside the sparse locations where we have measurements; i.e. it is a full scene reconstruction. In Fig. 7 we show a real example of gradient-based reconstruction for a scene with two planes at 27cm and 37cm. We captured a base scan in Fig. 7(a) of the scene and, using its depth gradients, captured a new set of measurements along edges Fig. 7(b). These were used with a widely available gradient reconstruction method [13] which reduced the corresponding RSME errors in Fig. 7(e) by a third.

V. LIMITATIONS AND CONCLUSION

Although we show only static scene reconstructions, our adaptive angular framework impacts any scanning TOF sensor.

Deblurring tradeoff: Given a minimum, required incident radiance at the photodetector, our sensor range Z and field-of-view Θ are inversely proportional, $Z^2 \propto \frac{1}{\tan(\frac{\Theta}{2})}$ (Fig. 8(I)). Our results have significant scope for improvement in measurement SNR of the reconstructions, and we will focus on better optical designs in the future.

System performance: In Fig. 8(II), we compare the ability to induce desired sample density on targets. For conventional sensors, as the density increases, the robot-target distance goes to zero. For our sensor design, a stand-off distance is possible since we can concentrate samples on the target.

Efficiency/power reduction applications: We will use energy efficient adaptive sensing for UAVs and other power constrained robots to place the samples on nearby obstacles and targets, accruing power savings.

REFERENCES

[1] Supreeth Achar, Joseph R Bartels, William L Whittaker, Kiriakos N Kutulakos, and Srinivasa G Narasimhan. Epipolar time-of-flight imaging. *ACM Transactions on Graphics (TOG)*, 36(4):37, 2017.

[2] Abhishek Badki, Orazio Gallo, Jan Kautz, and Pradeep Sen. Computational zoom: a framework for post-capture image composition. *ACM Transactions on Graphics (TOG)*, 36(4):46, 2017.

[3] Jacques M Beckers. Adaptive optics for astronomy: principles, performance, and applications. *Annual review of astronomy and astrophysics*, 31(1):13–62, 1993.

[4] Alvin T Campbell III and Donald S Fussell. Adaptive mesh generation for global diffuse illumination. In *ACM SIGGRAPH Computer Graphics*, volume 24, pages 155–164. ACM, 1990.

[5] Benjamin Charrow, Gregory Kahn, Sachin Patil, Sikang Liu, Ken Goldberg, Pieter Abbeel, Nathan Michael, and Vijay Kumar. Information-theoretic planning with trajectory optimization for dense 3d mapping. In *Robotics: Science and Systems*, 2015.

[6] Thomas Deselaers, Philippe Dreuw, and Hermann Ney. Pan, zoom, scantime-coherent, trained automatic video cropping. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[7] Rob Fergus, Antonio Torralba, and William T Freeman. Random lens imaging. 2006.

[8] Thomas P Flatley. Spacecube: A family of reconfigurable hybrid on-board science data processors. 2015.

[9] Joseph W Goodman et al. *Introduction to Fourier optics*, volume 2. McGraw-hill New York, 1968.

[10] Michael Gschwandtner, Roland Kwitt, Andreas Uhl, and Wolfgang Pree. Blensor: blender sensor simulation toolbox. In *International Symposium on Visual Computing*, pages 199–208. Springer, 2011.

[11] Mohit Gupta, Shree K Nayar, Matthias B Hullin, and Jaime Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Transactions on Graphics (ToG)*, 34(5):156, 2015.

[12] Ryan Halterman and Michael Bruch. Velodyne hdl-64e lidar for unmanned surface vehicle obstacle detection. Technical report, SPACE AND NAVAL WARFARE SYSTEMS CENTER SAN DIEGO CA, 2010.

[13] Matthew Harker and Paul OLeary. Regularized reconstruction of a surface from its measured gradient field. *Journal of Mathematical Imaging and Vision*, 51(1):46–70, 2015.

[14] Felix Heide, Matthias B Hullin, James Gregson, and Wolfgang Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics (ToG)*, 32(4):45, 2013.

[15] Geoffrey A Hollinger and Gaurav S Sukhatme. Sampling-based robotic information gathering algorithms. *The International Journal of Robotics Research*, 33(9):1271–1287, 2014.

[16] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, et al. Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 559–568. ACM, 2011.

[17] A. Jones, I. McDowall, H. Yamada, M. Bolas, and P. Debevec. Rendering for an interactive 360 degree light field display. In *SIGGRAPH*. ACM, 2007.

[18] Brian J Julian, Sertac Karaman, and Daniela Rus. On mutual information-based control of range sensing robots for mapping applications. *The International Journal of Robotics Research*, 33(10):1375–1392, 2014.

[19] Abhishek Kasturi, Veljko Milanovic, Bryan H Atwood, and James Yang. Uav-borne lidar with mems mirror-based scanning capability. In *Proc. SPIE*, volume 9832, page 98320M, 2016.

[20] Sanjeev J Koppal, Shuntaro Yamazaki, and Srinivasa G Narasimhan. Exploiting dlp illumination dithering for reconstruction and photography of high-speed scenes. *International journal of computer vision*, 96(1):125–144, 2012.

[21] Sanjeev J Koppal, Ioannis Gkioulekas, Travis Young, Hyunsung Park, Kenneth B Crozier, Geoffrey L Barrows, and Todd Zickler. Toward wide-angle microvision sensors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12):2982–2996, 2013.

[22] Krassimir T Krastev, Hendrikus WLAM Van Lierop, Herman MJ Soemers, Renatus Hendricus Maria Sanders, and Antonius Johannes Maria Nellissen. Mems scanning micromirror, September 3 2013. US Patent 8,526,089.

[23] Anat Levin, Rob Fergus, Frédo Durand, and William T Freeman. Image and depth from a conventional camera with a coded aperture. In *ACM Transactions on Graphics (TOG)*, volume 26, page 70. ACM, 2007.

[24] V Milanović, A Kasturi, N Siu, M Radojičić, and Y Su. memseye for optical 3d tracking and imaging applications. In *Solid-State Sensors, Actuators and Microsystems Conference (TRANSDUCERS), 2011 16th International*, pages 1895–1898. IEEE, 2011.

[25] Veljko Milanović, Abhishek Kasturi, James Yang, and Frank Hu. A fast single-pixel laser imager for vr/ar headset tracking. In *Proc. of SPIE Vol*, volume 10116, pages 101160E–1, 2017.

[26] Shree K Nayar, Vlad Branzoi, and Terry E Boult. Programmable imaging: Towards a flexible camera. *International Journal of Computer Vision*, 70(1):7–22, 2006.

[27] Diego Nehab, Pedro V Sander, and John R Isidoro. The real-time reprojection cache. In *ACM SIGGRAPH 2006 Sketches*, page 185. ACM, 2006.

[28] Ren Ng. Fourier slice photography. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 735–744. ACM, 2005.

[29] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B

Hullin, Wolfgang Heidrich, and Kiriakos N Kutulakos. Temporal frequency probing for 5d transient analysis of global light transport. *ACM Transactions on Graphics (ToG)*, 33(4):87, 2014.

[30] Matthew O'Toole, Supreeth Achar, Srinivasa G Narasimhan, and Kiriakos N Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM Transactions on Graphics (ToG)*, 34(4):35, 2015.

[31] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. *ACM Transactions on graphics (TOG)*, 22(3):313–318, 2003.

[32] Kurt E Petersen. Silicon torsional scanning mirror. *IBM Journal of Research and Development*, 24(5):631–637, 1980.

[33] Christopher V Poulton, Ami Yaacobi, David B Cole, Matthew J Byrd, Manan Raval, Diedrik Vermeulen, and Michael R Watts. Coherent solid-state lidar with silicon photonic optical phased arrays. *Optics letters*, 42(20): 4091–4094, 2017.

[34] Jose C Principe, Neil R Euliano, and W Curt Lefebvre. *Neural and adaptive systems: fundamentals through simulations*, volume 672. Wiley New York, 2000.

[35] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 179–188. ACM, 1998.

[36] Ramesh Raskar, Amit Agrawal, and Jack Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Transactions on Graphics (TOG)*, 25(3):795–804, 2006.

[37] Thilo Sandner, Claudia Baulig, Thomas Grasshoff, Michael Wildenhain, Markus Schwarzenberg, Hans-Georg Dahlmann, and Stefan Schwarzer. Hybrid assembled micro scanner array with large aperture and their system integration for a 3d tof laser camera. In *MOEMS and Miniaturized Systems XIV*, volume 9375, page 937505. International Society for Optics and Photonics, 2015.

[38] Barry L Stann, Jeff F Dammann, Mark Del Giorno, Charles DiBerardino, Mark M Giza, Michael A Powers, and Nenad Uzunovic. Integration and demonstration of mems–scanned ladar for robotic navigation. In *Proc. SPIE*, volume 9084, page 90840J, 2014.

[39] Demetri Terzopoulos and Manuela Vasilescu. Sampling and reconstruction with adaptive meshes. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*, pages 70–75. IEEE, 1991.

[40] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic robotics*. MIT press, 2005.

[41] Robert K Tyson. *Principles of adaptive optics*. CRC press, 2015.

[42] Andreas Velten, Di Wu, Belen Masia, Adrian Jarabo, Christopher Barsi, Chinmaya Joshi, Everett Lawson, Moungi Bawendi, Diego Gutierrez, and Ramesh Raskar. Imaging the propagation of light through scenes at picosecond resolution. *Communications of the ACM*, 59 (9):79–86, 2016.

[43] B Yang, L Zhou, X Zhang, S Koppal, and H Xie. A compact mems-based wide-angle optical scanner. In *Optical MEMS and Nanophotonics (OMN), 2017 International Conference on*, pages 1–2. IEEE, 2017.

[44] Hao Yang, Lei Xi, Sean Samuelson, Huikai Xie, Lily Yang, and Huabei Jiang. Handheld miniature probe integrating diffuse optical tomography with photoacoustic imaging through a mems scanning mirror. *Biomedical optics express*, 4(3):427–432, 2013.