Photon-Efficient Computational 3-D and Reflectivity Imaging With Single-Photon Detectors

Dongeek Shin, *Student Member, IEEE*, Ahmed Kirmani, *Student Member, IEEE*, Vivek K Goyal, *Fellow, IEEE*, and Jeffrey H. Shapiro, *Life Fellow, IEEE*

***Abstract*—Capturing depth and reflectivity images at low light levels from active illumination of a scene has wide-ranging applications. Conventionally, even with detectors sensitive to individual photons, hundreds of photon detections are needed at each pixel to mitigate Poisson noise. We develop a robust method for estimating depth and reflectivity using fixed dwell time per pixel and on the order of one detected photon per pixel averaged over the scene. Our computational image formation method combines physically accurate single-photon counting statistics with exploitation of the spatial correlations present in real-world reflectivity and 3-D structure. Experiments conducted in the presence of strong background light demonstrate that our method is able to accurately recover scene depth and reflectivity, while traditional imaging methods based on maximum likelihood (ML) estimation or approximations thereof lead to noisier images. For depth, performance compares favorably to signal-independent noise removal algorithms such as median filtering or block-matching and 3-D filtering (BM3D) applied to the pixelwise ML estimate; for reflectivity, performance is similar to signal-dependent noise removal algorithms such as Poisson nonlocal sparse PCA and BM3D with variance-stabilizing transformation. Our framework increases photon efficiency 100-fold over traditional processing and also improves, somewhat, upon first-photon imaging under a total acquisition time constraint in raster-scanned operation. Thus, our new imager will be useful for rapid, low-power, and noise-tolerant active optical imaging, and its fixed dwell time will facilitate parallelization through use of a detector array.**

***Index Terms*—3-D imaging, computational imaging, convex optimization, first-photon imaging, LIDAR, low-light imaging, Poisson noise, single-photon detection, time-of-flight imaging.**

I. Introduction

ACTIVE optical imaging methods measure properties of a scene using illumination provided by the system itself. Light detection and ranging (LIDAR) [1], also known as laser radar or LADAR, is a well-known example using periodically pulsed laser light. For each illuminated patch or pixel in an imaged scene, a LIDAR system builds a histogram of photon detection times, relative to the most recent pulse emission, over some number of pulses in a period called the *dwell time*. The time delay of this histogram, relative to the transmitted pulse's temporal profile, is related through the speed of light to the *depth* (equivalently, distance, range, or 3D structure) of the scene. The amplitude of this histogram is related to the *reflectivity* of the scene. For accurate depth and reflectivity estimation, the signal acquisition time must be long enough to collect the 10^2 to 10^3 photons per pixel (ppp) needed to generate a finely binned histogram for each pixel.

In this paper, we address the problem of achieving high photon efficiency in combined 3D and reflectivity imaging. We expound upon a framework, introduced in [2], that builds upon an approach initiated in [3], [4]. Like the first-photon imaging (FPI) method of [3], our computational imager distinguishes itself from other previous work by avoiding the formation of histograms and instead using probabilistic modeling at the level of individual detected photons. This physically accurate modeling of single-photon detection is combined with exploitation of the spatial correlations present in real-world scenes to achieve accurate 3D and reflectivity imaging from on the order of 1 detected ppp averaged over the scene, despite significant noise from background light and dark counts.

High photon efficiency is important when very little back-reflected light reaches the detector, as will be the case with low optical power relative to the imaging range [5]. More generally, increasing photon efficiency improves the trade-offs among optical power, imaging range, detector size, and imaging speed. The method introduced here uses a deterministic dwell time, which is both more convenient for raster scanning and amenable to parallelization through the use of a detector array. This ease of applicability comes with somewhat improved performance over FPI when compared at equal total acquisition times in raster-scanned operation.

A. Prior Work

1) Active Imaging Methods: Active 3D imaging systems differ in how they modulate their transmitted power, leading to a variety of trade-offs in accuracy, modulation frequency, optical power, and photon efficiency; see Fig. 1 for a qualitative summary. Temporal modulation enables absolute (unaliased) distance measurement by the time-of-flight (TOF) principle. Examples of TOF acquisition systems, ordered by increasing modulation bandwidth (decreasing pulse duration), include homodyne TOF cameras [6], pulsed TOF cameras [7],

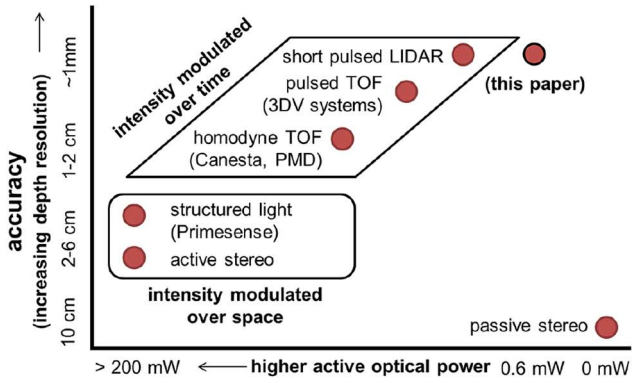
Manuscript received October 28, 2014; revised June 17, 2015; accepted June 30, 2015. Date of publication July 06, 2015; date of current version August 24, 2015. This work was supported in part by the U.S. National Science Foundation under Grant 1161413 and Grant 1422034, a Samsung Scholarship, and a Microsoft Ph.D. Fellowship. This work was presented in part at the IEEE International Conference on Image Processing, Paris, France, October 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alessandro Foi.

D. Shin, A. Kirmani, and J. H. Shapiro are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: dshin@mit.edu; akirmani@mit.edu; jhs@mit.edu).

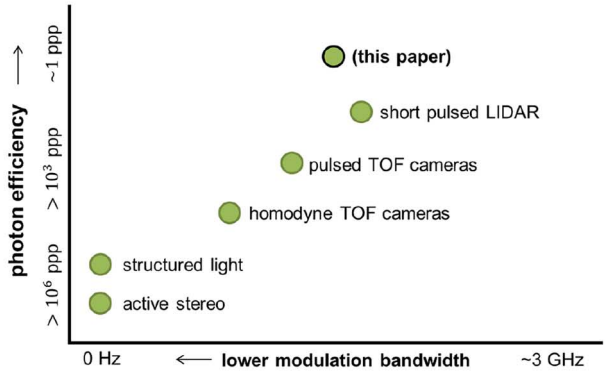
V. K. Goyal is with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA (e-mail: v.goyal@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCI.2015.2453093



(a) Accuracy vs. power trade-offs.



(b) Photon efficiency vs. modulation bandwidth trade-offs.

Fig. 1. Qualitative comparison of state-of-the-art active optical 3D sensing technologies. Photon efficiency is defined as photons per pixel (ppp) necessary for centimeter-accurate depth imaging.

and picosecond laser radar systems [8]. Spatial modulation techniques include structured light [9] and active stereo imaging [10]. These spatial-modulation techniques have low photon efficiencies because they use an always-on optical source, whereas pulsed-TOF systems have higher photon efficiencies because they use sources that are on only for short intervals. Additionally, the systems using temporal modulation have better absolute-distance accuracy than those using spatial modulation. The advantage of spatial modulation tends to be cheaper sensing hardware, since high-speed sampling is not required.

The most photon-efficient TOF imagers—those requiring the fewest photons for accurate imaging—use single-photon avalanche diode (SPAD) detectors [11]. Earlier efforts in SPAD-based 3D imaging from on the order of 1 detected ppp are reported in [12]–[14]. The framework presented here improves upon these works in part due to the use of estimated reflectivity. This translates to SPAD-based imagers with lower optical power and lower system bandwidth without sacrificing image quality. There also has been significant recent interest in compressive methods for 3D imaging, with [15]–[17] and without [18] single-photon detection. While compressive methods may reduce some measures of acquisition cost, they do not generally improve photon efficiency.

2) Optoelectronic Techniques for Low Light Levels: In low-light scenarios, a variety of optoelectronic techniques are employed to improve robustness. Active imagers use lasers with narrow spectral bandwidths and spectral filters to suppress background light and minimize the Poisson noise it creates. However, optical filtering alone cannot completely eliminate background light, and it also causes signal attenuation. Range-gated imaging [19] is another common technique, but this method requires a priori knowledge of object location. Furthermore, a SPAD may be replaced with a superconducting nanowire single-photon detector (SNSPD) [20], which is much faster, has lower timing jitter, and has lower dark-count rate than a SPAD. However, SNSPDs have much smaller active areas and hence have narrower fields of view than SPAD-based systems with the same optics.

3) Image Denoising: For depth imaging using SPAD data, it is typical to first form an image from a pixel-by-pixel estimate of scene depth using a time-inhomogeneous Poisson process model for photon detection times and then to apply an image denoising method that exploits the scene's spatial correlations. As discussed in Section IV-B, even for an individual pixel, maximum likelihood (ML) estimation is considered prohibitively complex and thus is typically replaced by a log-matched filter, which is an approximation to it. In this two-step approach of pixelwise estimation and denoising, one commonly assumes a Gaussian noise model for the output of the first step. This is empirically justified for high light levels [21] and also justified by the asymptotic normality of ML estimates with large numbers of data samples. However, at low light levels with significant background light, performing denoising well is more challenging due to the resulting high-variance uniform noise on the observed photon arrival times. In Section VI, we compare our technique with a state-of-the-art denoising method that uses block matching. The superior performance of our method is due in part to classification of photon detection events as being due to signal (backscattered light) or noise (unwanted background light and dark counts) prior to any depth image formation; classical pixelwise image formation yields an extremely challenging denoising problem. Note that denoising of reflectivity images is a better-developed field than denoising of depth images, but existing methods are not designed for the very low light levels and regularly-textured natural scenes considered in this paper.

4) First-Photon Imaging: First-photon imaging (FPI) [3] is a method that forms 3D and reflectivity images using only the first detected photon at every pixel in a raster-scanned scene. FPI combines accurate first-photon detection statistics with the spatial correlations existing in natural scenes to achieve robust low-light imaging. The use of the first detection time in FPI, however, makes the dwell time at each pixel a random variable. Thus, FPI does not extend naturally from raster-scanned data collection to the use of SPAD arrays—since simultaneous measurement implies equal dwell times—thus precluding the dramatic speedup in image acquisition that such arrays enable.

In this paper, we develop models and methods analogous to FPI that apply when there is a fixed dwell time at each pixel. In the experimental configuration depicted in Fig. 2, we demonstrate that the performance of the new method is similar to or slightly better than FPI when compared for equal total acquisition time in raster-scanned operation. Furthermore, with an M-fold increase in laser power and an M-element SPAD array,

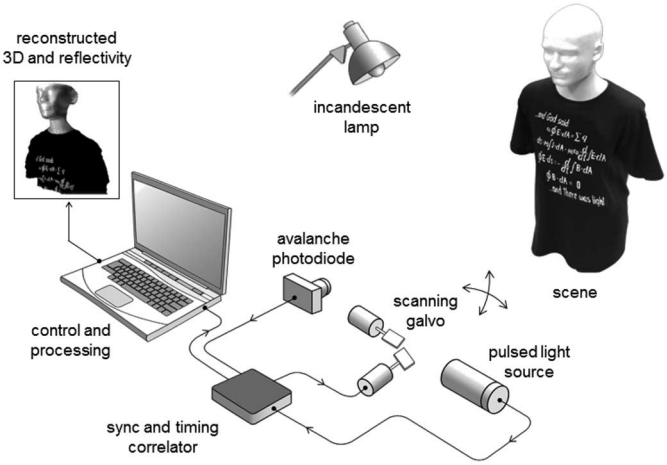


Fig. 2. Experimental imaging setup used with random dwell time in [3] and with constant dwell time here. A pulsed light source illuminates the scene in a raster scan pattern. The backscattered light is collected by a time-resolved single-photon detector. Each spatial location is illuminated with exactly N light pulses (fixed dwell time). An incandescent lamp injects background light which corrupts the information-bearing signal. The photon detection times and the total photon count are recorded at every image pixel. This dataset is used to estimate the 3D structure and reflectivity. The setup is analogous to having a floodlight illumination source and an array of single-photon counting detectors operating at a fixed dwell time.

our fixed dwell-time framework can provide this same robust imaging M times faster than a single-detector raster-scanned system.

B. Main Contributions

1) Modeling: We introduce a physically accurate model for the signal produced by a SPAD under low-light conditions that incorporates an arbitrary illumination pulse shape, background (ambient) light contribution, dark counts, and the inhomogeneous Poisson process characteristics (shot noise from the quantum nature of light) given a fixed dwell time. The same model for a single illumination pulse was used in [3] (with limited explanation); while [3] used a random number of pulses, the analysis in this paper is for a fixed number of pulses.

2) Algorithmic: We provide a method for computational reconstruction of depth and reflectivity from noisy photon-detection data. Our technique combines a shot-noise model for single-photon detection with simple means to exploit the high degree of spatial correlation present in real-world scenes. The modularity of the technique—combining spatial regularization for reflectivity, classification of detections as due to signal or noise, and spatial regularization for depth—makes it amenable to the generation of algorithmic variations to exploit more sophisticated spatial correlation models.

3) Experimental: We experimentally demonstrate that our proposed 3D imager's photon efficiency is more than 100 times higher than that of the conventional log-matched filter, which is a well-known proxy for pixelwise ML estimation. We also show that our 3D imager achieves sub-pulse-width depth resolution under short acquisition times, in which 54% of the pixels have missing data (no photon detections), and at high background levels, when any given photon detection has approximately probability 0.5 of originating from ambient light.

C. Outline

The remainder of the paper is organized as follows. Section II introduces the LIDAR-like imaging configuration that we consider. The key probabilistic models for the measured data are derived in Section III. These models are related to conventional image formation in Section IV, and they are the basis for the novel image formation method in Section V. Section VI presents experimental results for the novel method, and Section VII provides additional discussion and conclusions. An appendix presents performance bounds for pixelwise estimators based on our modeling.

The methods detailed in this paper were presented in preliminary, abbreviated form in [2]. The present manuscript provides additional context (Section I-A), details on derivations (Section III), performance bounds (Appendix), and experimental results that do not appear in [2] (Section VI). In particular, comparisons to Poisson non-local sparse PCA (NLSPCA) [22] and block-matching and 3D filtering (BM3D) [23]—with variance-stabilizing transform (VST) for reflectivity [24]—replace the use of less-sophisticated bilateral filtering [25] in the preliminary work.

II. Imaging Setup

Fig. 3 depicts the signal acquisition model underlying our imager. We aim to form reflectivity image $\boldsymbol{\alpha} \in \mathbb{R}_+^{n\times n}$ and depth image $\mathbf{z} \in \mathbb{R}_+^{n\times n}$ of the scene. We index the scene pixels by $(i,j) \in \{1,\,2,\,\ldots,\,n\}^2$. The distance to pixel (i,j) is $z_{i,j} \in [0,\,z_{\max})$, where $z_{\max}$ is a known maximum scene depth, and its reflectivity, $\alpha_{i,j} \geq 0$, includes the effects of radial fall-off, view angle, and material properties.

A. Illumination

We use a periodically pulsed laser to illuminate the scene in a raster-scanned manner. The repetition period is T_r and the waveform of a single pulse is denoted by $s(t)$. Physically, $s(t)$ is the photon-flux waveform of the pulse emitted at $t = 0$ measured in counts/sec (cps). To avoid distance aliasing, we assume $T_r > 2z_{\max}/c$, where c is the speed of light, and the support of $s(t)$ is contained in $[0, T_r]$. With conventional processing, the root mean-square (RMS) pulse width T_p governs the achievable depth resolution in the absence of background light [26]. As is typically done in depth imaging, we assume that $T_p \ll 2z_{\max}/c$.

B. Detection

A SPAD detector provides time-resolved single-photon detections [11], called *clicks*. Its quantum efficiency η is the fraction of photons passing through the pre-detection optical filter that are detected. Each detected photon is time stamped within a time bin of duration measuring a few picoseconds. We will assume that this quantization of detection times is much shorter than T_p and thus negligible; i.e., we treat the detection times as continuous-valued variables.

A SPAD detector is not *number-resolving*, meaning that it reports at most one click in a short period of time. This is because SPAD detectors have a *reset time* or *dead time* after

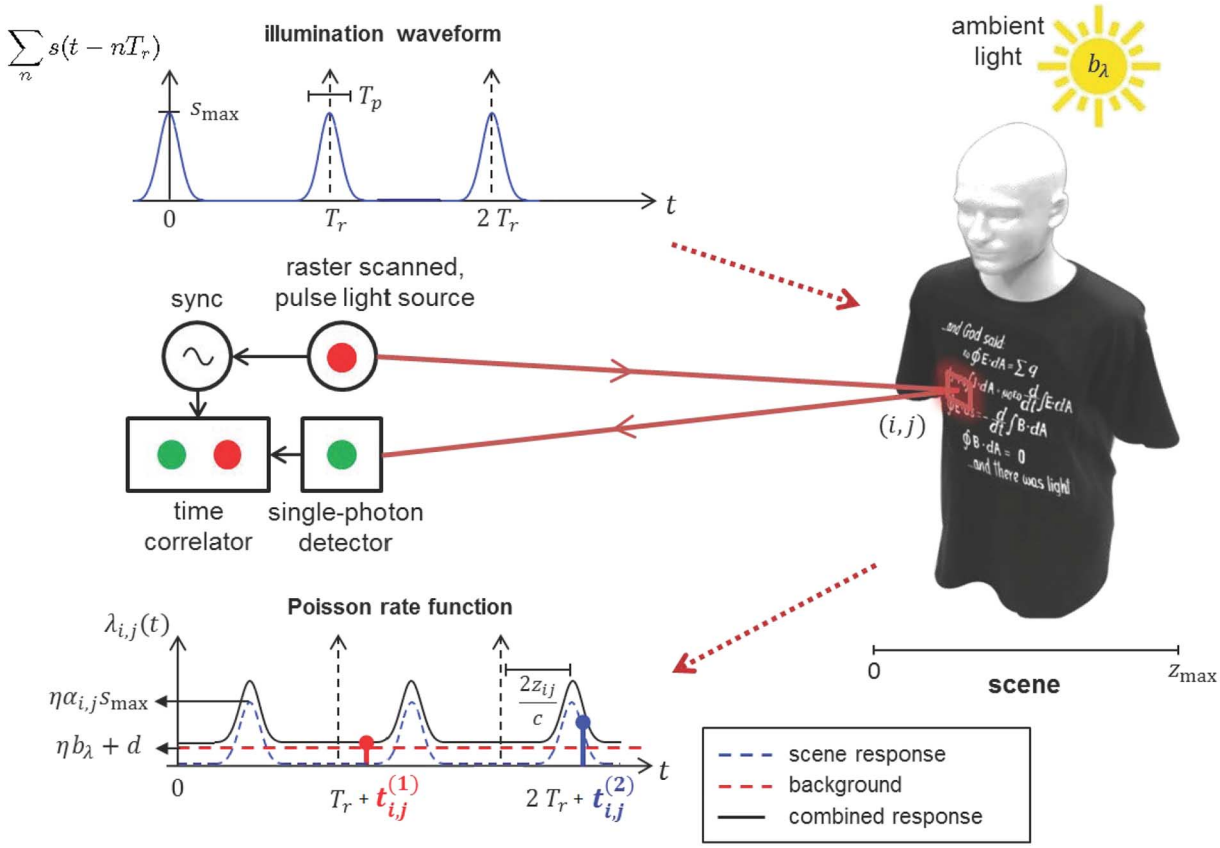


Fig. 3. Summary of observation model. Rate function of inhomogeneous Poisson process combining desired scene response and noise sources is shown. Here, $N = 3$ and $k_{i,j} = 2$. A background count (red) occurred after the second pulse was transmitted, and a signal count (blue) occurred after the third pulse was transmitted.

each click, during which there is no sensitivity to incident light. Here, we consider low-flux imaging where the probability of multiple clicks within one repetition period of duration T_r would be negligible even without reset time.

C. Data Acquisition

Each pixel (i, j) is illuminated with a sequence of N laser pulses. Our methods are intended for values of N on the order of 10 to 100. The total dwell time is thus $T_a = NT_r$. To simulate significant noise from ambient light and dark counts, we also shine background light, with photon flux b_λ at the operating optical wavelength λ, onto the detector. For each pixel, we record the total number of photon detections $k_{i,j}$, along with their detection times $\{t_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}}$, where the latter are measured relative to the immediately-preceding transmitted pulse.

III. Probabilistic Measurement Model

Illuminating pixel (i, j) with the pulse $s(t)$ results in backreflected light with photon flux

$$r_{i,j}(t) = \alpha_{i,j} s(t - 2z_{i,j}/c) + b_\lambda$$

at the detector. The measurement of photon flux is through photon detections, and carefully modeling the relationships between the measured quantities and the reflectivity and depth variables is central to our imaging method.

A. Poisson Statistics

Ignoring SPAD reset time, because of the low-flux condition, the photon detections produced by the SPAD in response to the backreflected light from transmission of $s(t)$ constitute an inhomogeneous Poisson process [27] with time-varying rate function $\eta r_{i,j}(t)$. To these photon detections we must add the detector dark counts, which come from an independent homogeneous Poisson process with rate d. Lumping the dark counts together with the background-generated clicks yields the observation process at the SPAD's output, viz., as shown in Fig. 3, an inhomogeneous Poisson process with rate function

$$\begin{aligned}\lambda_{i,j}(t) &= \eta r_{i,j}(t) + d \\ &= \eta\alpha_{i,j} s(t - 2z_{i,j}/c) + (\eta b_\lambda + d),\end{aligned} \tag{1}$$

when only a single pulse is transmitted. This means that the probability mass function (PMF) for the number of arrivals $M(t_1, t_2)$ in time interval $(t_1, t_2]$ is

$$p(k) = \frac{m(t_1,t_2)^k e^{-m(t_1,t_2)}}{k!}, \quad k = 0, 1, \ldots, \tag{2a}$$

where

$$m(t_1, t_2) = \int_{t_1}^{t_2} \lambda_{i,j}(\tau)\, d\tau. \tag{2b}$$

Fig. 3 shows the rate function $\lambda_{i,j}(t)$ for the periodic transmission.

Define $S = \int_0^{T_r} s(t)\, dt$ and $B = (\eta b_\lambda + d)T_r$ as the total signal and background count per pulse-repetition period, where we have used—and will use in all that follows—background counts to include dark counts as well as counts arising from ambient light. We assume that B is known, because it is straightforward to measure it before we begin data acquisition. The derivations to follow assume $\eta\alpha_{i,j}S + B \ll 1$, meaning that the photon-flux per pixel per pulse-repetition period is much less than 1, as would be the case in low-light imaging where an imager's photon efficiency is paramount. This low flux per pulse-repetition period necessitates imaging over many pulse-repetition periods, i.e., at least a moderate value of N.

B. Distributions of Numbers of Detected Photons

For the moment, imagine the use of a number-resolving detector rather than a SPAD; i.e., disregard detector reset time. Using the rate function given in (1), the PMF for the number of photons detected in response to a single illumination pulse is given by (2a), where

$$m(0, T_r) = \int_0^{T_r} \lambda_{i,j}(\tau)\, d\tau = \eta\alpha_{i,j}S + B \tag{3}$$

is the mean number of detections. In particular, the probability of zero detections is thus

$$P_0(\alpha_{i,j}) = \exp[-(\eta\alpha_{i,j}S + B)]. \tag{4}$$

The low-flux condition ensures that the probability of multiple photon detections per repetition interval is negligible relative to the probability of zero or one detection.

Returning to imaging with a SPAD, we either have zero detections with the probability given in (4) or one detection with the complementary probability. Then, because detections across the N illumination pulses are statistically independent, the number of detected photons $K_{i,j}$ is binomially distributed with PMF

$$\begin{aligned}&\Pr\left[K_{i,j} = k_{i,j}; \alpha_{i,j}\right]\\ &\quad = \binom{N}{k_{i,j}} P_0(\alpha_{i,j})^{N-k_{i,j}} \left[1 - P_0(\alpha_{i,j})\right]^{k_{i,j}}, \end{aligned}\tag{5}$$

for $k_{i,j} = 0, 1, \ldots, N$.

Unlike previous works, our imaging method uses the PMF in (5) explicitly. Note that for large N with $P_0(\alpha_{i,j})$ held constant, the binomial PMF is well-approximated as Gaussian; this justifies the traditional use of an additive Gaussian noise model when the light level is high. Similarly, in the ultimate low-flux limit in which $\eta\alpha_{i,j}S + B \to 0^+$ with $N \to \infty$ such that $C(\alpha_{i,j}) = N\{1 - \exp[-(\eta\alpha_{i,j}S + B)]\}$ is held constant, $K_{i,j}$ converges to a Poisson random variable [21] with PMF

$$\Pr[K_{i,j} = k_{i,j};\ \alpha_{i,j}] = \frac{C(\alpha_{i,j})^{k_{i,j}}}{k_{i,j}!} \exp[-C(\alpha_{i,j})]; \tag{6}$$

this model is often used for low light levels.

C. Distributions of Photon Detection Times

In our imaging system, the SPAD detects at most one photon per pulse-repetition period. The time of this detection is equivalent to the first detection time of a system not limited by the reset time of a SPAD detector. Thus, we are interested in the distribution of the first arrival time in an inhomogeneous Poisson process with rate function given in (1), given that there is at least one arrival.

Let U denote the first arrival time. For any $u \in (0, T_r)$, the conditional cumulative distribution function of U given that there is at least one arrival in $(0, T_r]$ satisfies

$$\begin{aligned} F(u) &= \Pr[U \le u | M(0, T_r) \ge 1]\\ &= \Pr[M(0, u) \ge 1 | M(0, T_r) \ge 1]\\ &\overset{(a)}{=} \frac{\Pr[\{M(0, u) \ge 1\} \cap \{M(0, T_r) \ge 1\}]}{\Pr[M(0, T_r) \ge 1]}\\ &\overset{(b)}{=} \frac{\Pr[M(0, u) \ge 1]}{\Pr[M(0, T_r) \ge 1]}\\ &= \frac{1 - \Pr[M(0, u) = 0]}{1 - \Pr[M(0, T_r) = 0]}\\ &\overset{(c)}{=} \frac{1 - \exp\left[-\int_0^u \lambda_{i,j}(\tau)\, d\tau\right]}{1 - P_0(\alpha_{i,j})}, \end{aligned}\tag{7}$$

where (a) follows from the definition of conditioning; (b) from $u \in (0, T_r)$; and (c) from (2). Differentiating (7) with respect to u gives the conditional probability density function (PDF) of U to be

$$f_U(u) = \frac{\lambda_{i,j}(u) \exp\left[-\int_0^u \lambda_{i,j}(\tau)\, d\tau\right]}{1 - P_0(\alpha_{i,j})}. \tag{8}$$

Returning to our imaging system, by using the low-flux assumption that $\eta\alpha_{i,j}S + B \ll 1$, we incur very little error by neglecting the exponential factor in (8). Thus, for any illumination pulse at pixel (i, j) for which there is a detected photon, the PDF of the detection time $T_{i,j}$ is modeled as

$$\begin{aligned} f_{T_{i,j}}(t_{i,j};\ \alpha_{i,j}, z_{i,j}) &= \frac{\lambda_{i,j}(t_{i,j})}{1 - P_0(\alpha_{i,j})}\\ &= \frac{\eta\alpha_{i,j}s(t_{i,j} - 2z_{i,j}/c) + B/T_r}{1 - P_0(\alpha_{i,j})}\\ &\overset{(a)}{\approx} \frac{\eta\alpha_{i,j}s(t_{i,j} - 2z_{i,j}/c) + B/T_r}{\eta\alpha_{i,j}S + B}\\ &= \frac{\eta\alpha_{i,j}S}{\eta\alpha_{i,j}S + B}\left(\frac{s(t_{i,j} - 2z_{i,j}/c)}{S}\right)\\ &\quad + \frac{B}{\eta\alpha_{i,j}S + B}\left(\frac{1}{T_r}\right), \end{aligned}\tag{9}$$

where (a) follows from $1 - e^{-x} \approx x$ for $0 < x \ll 1$. An alternate derivation of this PDF that includes the time resolution of the detector is given in [28].

A detection could be a signal count or a background count. The detection statistics result from the merging of the Poisson processes corresponding to these sources. Under our low-flux assumption, the detection time for a signal count from a single pulse-repetition interval is characterized by the normalized

time-shifted pulse shape. On the other hand, the detection time for a background count in that interval is uniformly distributed on $[0, T_r)$; this is a standard property for the arrival time in a homogenous Poisson process conditioned on having exactly one arrival. These two distributions are apparent in the PDF in (9); it is a mixture distribution, with mixture weights

$$\Pr[\text{Detection at } (i,j) \text{ is signal}] = \frac{\eta\alpha_{i,j}S}{\eta\alpha_{i,j}S + B},$$
$$\Pr[\text{ Detection at } (i,j) \text{ is noise}] = \frac{B}{\eta\alpha_{i,j}S + B}.$$

IV. Conventional Image Formation

A. Pixelwise Reflectivity Estimation

Given the total observed photon count $k_{i,j}$ at pixel (i,j), the constrained ML (CML) reflectivity estimate is

$$\hat{\alpha}_{i,j}^{\text{CML}} = \underset{\alpha_{i,j}\geq 0}{\arg\max}\ \Pr[K_{i,j} = k_{i,j}; \alpha_{i,j}]$$
$$= \max\left\{\frac{1}{\eta S}\left[\log\left(\frac{N}{N - k_{i,j}}\right) - B\right], 0\right\},$$

where we have used (5), log is the natural logarithm, and the constraint is simply the nonnegativity of reflectivity. Traditionally, the normalized photon-count value is used as an estimate for reflectivity [29],

$$\tilde{\alpha}_{i,j} = \frac{k_{i,j}}{N\eta S}. \tag{10}$$

Because the probability mass function of the photon count is

$$\Pr[K_{i,j} = k_{i,j}; \alpha_{i,j}] = \frac{\exp[-N(\eta\alpha_{i,j}S + B)](N(\eta\alpha_{i,j}S + B))^{k_{i,j}}}{k_{i,j}!} \tag{11}$$

under the Poisson approximation to the binomial distribution in (6), we note that the normalized photon-count estimate $\tilde{\alpha}$ is equal to the ML estimate under the Poisson approximation to the binomial distribution when $B = 0$.

B. Pixelwise Depth Estimation

Assuming that the photon detection-time dataset $\{t_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}}$ at pixel (i,j) is non-empty, the CML depth estimate is given in principle by

$$\hat{z}_{i,j}^{\text{CML}} = \underset{z_{i,j}\in[0,z_{\max})}{\arg\max} \prod_{\ell=1}^{k_{i,j}} f_{T_{i,j}}(t_{i,j}^{(\ell)};\ \alpha_{i,j}, z_{i,j})$$
$$= \underset{z_{i,j}\in[0,z_{\max})}{\arg\max} \sum_{\ell=1}^{k_{i,j}} \log\left[\eta\alpha_{i,j} s\left(t_{i,j}^{(\ell)} - \frac{2z_{i,j}}{c}\right) + \frac{B}{T_r}\right], \tag{12}$$

where we have used (9). However, this estimate rarely used in practice. If $B > 0$, this optimization problem is non-convex. Moreover, when $B > 0$, this CML estimation requires knowledge of the true reflectivity $\alpha_{i,j}$, which is not typically available. Thus, to solve (12) for an arbitrary pulse waveform $s(t)$, the optimization depends on discretization of $z_{i,j}$ and of $\alpha_{i,j}$ and a subsequent search; since $T_p \ll T_r$, the search may be prohibitively expensive to achieve depth resolution on the order of $T_p c$.

For reduced complexity, the log-matched filter [27] is instead traditionally used for estimating depth from $k_{i,j} \geq 1$ photon detections:

$$\tilde{z}_{i,j} = \underset{z_{i,j}\in[0,z_{\max})}{\arg\max} \sum_{\ell=1}^{k_{i,j}} \log\left[s\left(t_{i,j}^{(\ell)} - 2z_{i,j}/c\right)\right]. \tag{13}$$

The log-matched filter solution is equal to the CML estimate when $B = 0$, since then the scaling by $\eta\alpha_{i,j}$ is irrelevant to the location of the maximum with respect to $z_{i,j}$. For a Gaussian pulse waveform s centered at time zero, the log-matched filter estimate $\tilde{z}_{i,j}$ simplifies further to the sample mean of $\{t_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}}$. This simpler estimate is also commonly employed [30].

V. Novel Image Formation

In the limit of large sample size or high signal-to-noise ratio (SNR), the ML estimate converges to the true parameter value [31]. However, when the data is limited or SNR is low—such as in our problem—pixelwise ML solutions yield inaccurate estimates. We compare our 3D imaging method with the baseline normalized-count reflectivity estimate $\tilde{\alpha}_{i,j}$ from (10) and the log-matched filter depth estimate $\tilde{z}_{i,j}$ from (13), which are ML estimates asymptotically in the signal count. Along with using the single-photon detection statistics, we exploit the spatial correlations present in real-world scenes by regularizing these estimators. Our approach provides significant improvements over pixelwise estimators as well as over denoising techniques that may better exploit scene structure but assume additive Gaussian noise or implicitly assume at least moderate light levels. Our computational image formation proceeds in three steps.

Step 1: Reflectivity estimation. Using (5), the negative log-likelihood of scene reflectivity $\alpha_{i,j}$ given count data $k_{i,j}$ is

$$\mathcal{L}_\alpha(\alpha_{i,j}; k_{i,j}) = (N - k_{i,j})\eta S\alpha_{i,j} - k_{i,j}\log\left\{1 - \exp[-(\eta\alpha_{i,j}S + B)]\right\}, \tag{14}$$

after constants independent of $\alpha_{i,j}$ are dropped. Since $\mathcal{L}_\alpha(\alpha_{i,j}; k_{i,j})$ is a strictly convex function of $\alpha_{i,j}$ on $[0,\infty)$, it is amenable to global minimization using convex optimization, with or without the inclusion of sparsity-based regularization [32]. The penalized ML (PML) estimate for scene reflectivity is obtained from noisy data $\{k_{i,j}\}_{i,j=1}^n$ by solving the following convex program:

$$\hat{\boldsymbol{\alpha}}^{\text{PML}} = \underset{\alpha:\alpha_{i,j}\geq 0}{\arg\min} \sum_{i=1}^n \sum_{j=1}^n \mathcal{L}_\alpha(\alpha_{i,j}; k_{i,j}) + \beta_\alpha\ \text{pen}_\alpha(\boldsymbol{\alpha}), \tag{15}$$

where $\text{pen}_\alpha(\cdot)$ is a convex function that penalizes the non-smoothness of the reflectivity estimate, and β_α controls the degree of penalization.

Step 2: Rejection of background detections. Direct application of a similar regularized-ML approach to depth estimation using time-of-detection data is infeasible. This is because the background contribution to the likelihood function creates a non-convex cost function with locally optimal solutions that are far from the global optimum. Hence, before estimating depth, a second processing step attempts to identify and censor the detections that are due to background.

Background counts do not contain any scene depth information. Their detection times are mutually independent over spatial locations with variance $T_r^2/12$. In contrast, since light pulses have duration $T_p \ll T_r$ and depths $z_{i,j}$ are correlated over spatial locations, the detection times of signal counts have conditional variances, given data from neighboring positions, that are much lower than $T_r^2/12$. Based on this key observation, our method to censor a noisy detection at (i,j) is as follows:

1) Compute the rank-ordered mean (ROM) $t_{i,j}^{\text{ROM}}$ for each pixel, which is the median value of the detection times at the 8 neighboring pixels of (i,j) [33]. If $t_{i,j}^{\text{ROM}}$ cannot be computed due to missing data (there are no detections for all of the neighboring pixels), set $t_{i,j}^{\text{ROM}} = \infty$.
2) Form the set of uncensored detections (those presumed to be signal detections), as follows:

$$U_{i,j} = \left\{ \ell \in \{1,\dots,k_{i,j}\} : |t_{i,j}^{(\ell)} - t_{i,j}^{\text{ROM}}| < 2T_p \left(\frac{B}{\eta \hat{\alpha}_{i,j}^{\text{PML}} S + B} \right) \right\}.$$

If $k_{i,j} = 0$, then set $U_{i,j} = \emptyset$.[1]

The ROM filter is different from the standard median filter because it uses signal values from only the neighboring pixels and not the reference pixel. It is demonstrated in [33] that the ROM filtered image t^{ROM} is a good approximation of the true image when the true image is corrupted by high-variance impulse noise at every pixel. In our imaging setup, the background photon detections are uniformly distributed with high variance. Also, the variances of signal photon detections depend on the duration of the transmitted pulse. Thus, the quality of ROM approximation at pixel (i,j) deteriorates as the probability of detecting a background photon $B/(\eta\alpha_{i,j}S + B)$ increases or the RMS pulse width T_p increases. Because the condition for censoring photon detections must be relaxed for an unreliable ROM estimate at every pixel, we set the censoring threshold parameter to be linearly dependent on both the RMS pulse width and our estimate of the background detection probability. We have found that removing dependence on $\hat{\alpha}_{i,j}^{\text{PML}}$ from the censoring rule results in significantly worse performance; this link between estimation of reflectivity and depth is a feature common to this work and FPI but not seen in earlier methods for photon-efficient imaging.

Step 3: Depth Estimation. Neglecting the second term of (9) because background detections have been rejected, the negative log-likelihood function of depth $z_{i,j}$, given uncensored data $\{t_{i,j}^{(\ell)}\}_{\ell \in U_{i,j}}$ is

$$\mathcal{L}_z\left(z_{i,j};\ \{t_{i,j}^{(\ell)}\}_{\ell\in U_{i,j}}\right) = -\sum_{\ell \in U_{i,j}} \log\left[s\left(t_{i,j}^{(\ell)} - 2z_{i,j}/c\right)\right],$$

after constants independent of $z_{i,j}$ are dropped. Note the simplicity compared to the CML estimate (12) and the similarity to the log-matched filter (13) (with detections presumed to be due to background omitted). If $U_{i,j} = \emptyset$, then set $\mathcal{L}_z(z_{i,j}; \{t_{i,j}^{(\ell)}\}_{\ell\in U_{i,j}}) = 0$, so that it has no contribution to the scene's negative log-likelihood cost function.

Our modeling to this point allows the illumination pulse shape $s(t)$ to be arbitrary, but properties of $s(t)$ determine the difficulty of computing ML or penalized ML estimates. Many practical pulse shapes are well approximated using log-concave functions (such as Gaussian), where $s(t) \propto \exp[-v(t)]$ and $v(t)$ is a convex function in t. Then,

$$\mathcal{L}_z(z_{i,j}; \{t_{i,j}^{(\ell)}\}_{\ell\in U_{i,j}}) = \sum_{\ell\in U_{i,j}} v(t_{i,j}^{(\ell)} - 2z_{i,j}/c)$$

is a convex function in $z_{i,j}$. Our penalized ML estimate for the scene depth image is thus obtained using uncensored data and solving the following convex optimization problem:

$$\hat{\mathbf{z}}^{\text{PML}} = \underset{z:z_{i,j}\in[0,z_{\max})}{\arg\min} \sum_{i=1}^{n}\sum_{j=1}^{n} \mathcal{L}_z\left(z_{i,j}; \{t_{i,j}^{(\ell)}\}_{\ell\in U_{i,j}}\right) + \beta_z \text{pen}_z(\mathbf{z}), \tag{16}$$

where $\text{pen}_z(\cdot)$ is a convex function that penalizes non-smoothness of the depth estimate and $\beta_z > 0$ controls the degree of penalization.

Choices of Regularizers. The regularizers $\text{pen}_\alpha(\cdot)$ and $\text{pen}_z(\cdot)$ should be chosen based on characteristics of the scenes of interest. Our framework is not tied to a particular choice of the regularizers, though convex regularizers are needed for the optimization problems (15) and (16) to be convex. Our experiments in the next section use the total variation semi-norm [34], [35] for both $\text{pen}_\alpha(\cdot)$ and $\text{pen}_z(\cdot)$. In a Bayesian setting, using regularizers proportional to the negative log prior makes the PML estimators equivalent to maximum a posteriori (MAP) estimators.

VI. Experimental Results

To test the performance of our proposed 3D structure and reflectivity imaging method, we used the photon arrival dataset collected by D. Venkatraman for [3], which is available from [36]. The experimental setup used to collect data is shown in Fig. 2. A pulsed laser diode with pulse width $T_p = 270$ ps and repetition period $T_r = 100$ ns was used as the illumination source. A two-axis galvonometer was used to raster scan 1000×1000 pixels, with the goal of angular uniformity along both axes. A lensless SPAD detector with quantum efficiency $\eta = 0.35$ was used for detection. The background light level was set such that B equaled the scene-averaged value of $\eta\alpha_{i,j}S$. Further details of the experimental setup are given in [3],

[1]Through the definition of $t_{i,j}^{\text{ROM}}$, we also have $U_{i,j} = \emptyset$ when there are no detections for all of the neighboring pixels. In this case, it is likely from correlations in reflectivity that the reflectivity at (i,j) is very low; hence, it is likely that detections at (i,j) are due to noise.

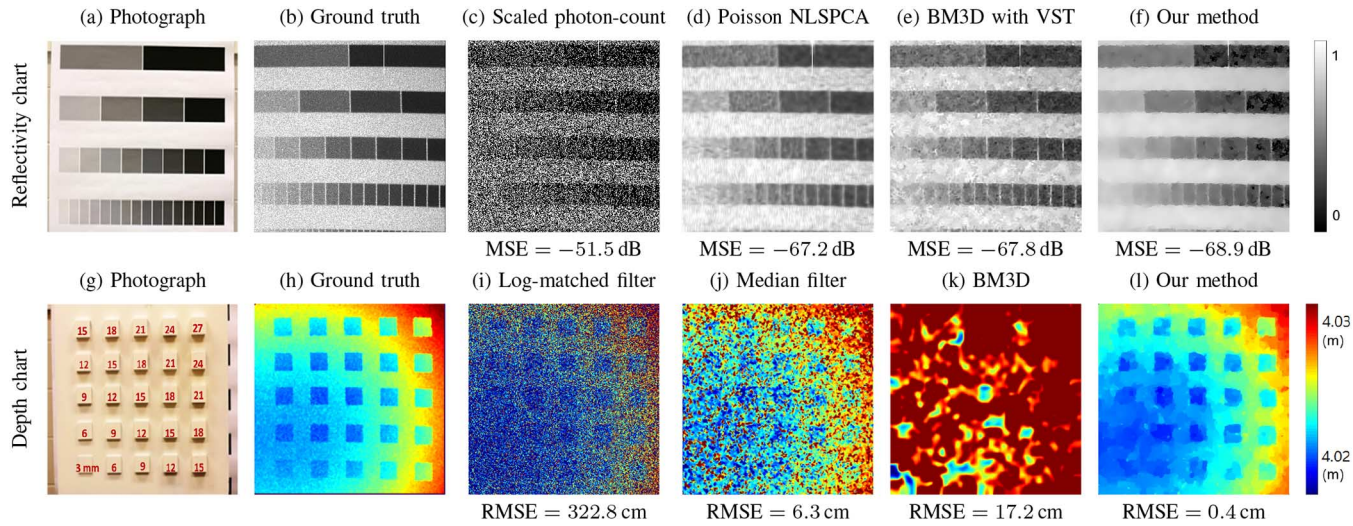


Fig. 4. Resolution test experiments. Reflectivity chart imaging (top) was done using $T_a = 30$ μs and had a mean count per pixel of 0.48. They were scaled such that the relevant reflectivity features are captured in the interval [0,1]. Depth chart imaging (bottom) was done using $T_a = 6.2$ μs and had a mean count per pixel of 1.1 with 33% of the pixels having missing data, i.e., no detections.

[4] and its supplementary material [37]. Because ideal raster scanning with a fixed dwell time is equivalent to using a flood-light illumination source and a detector array, our experimental results are indicative of what can be accomplished in real-time imaging scenarios using SPAD arrays.

A. *Reflectivity Resolution Test*

Reflectivity resolution was evaluated using the linear grayscale reflectivity chart shown in Fig. 4(a). To play the role of a reference (ground-truth) image, we formed the pixelwise ML estimate shown in Fig. 4(b) from about 1000 photon detections per pixel collected with a dwell time of about 60 ms. Evaluation of our method used a much shorter dwell time of $T_a = 30$ μs, which resulted in mean photon count per pixel of 0.48. Fig. 4(f) shows that our image formation method (Equation (15) with the total variation semi-norm for $\text{pen}_\alpha(\cdot)$) resolves the 16 gray levels present in the chart. The conventional normalized photon-count estimate (Equation (10)) yields the very noisy image shown in Fig. 4(c). Since $\tilde{\alpha}_{i,j}$ is a scaled Poisson random variable, this image can presumably be improved using existing state-of-the-art image denoising algorithms suitable for Poisson observations. Fig. 4(d) shows that the result of applying Poisson non-local sparse PCA (NLSPCA) [22]; the parameters were set using the suggestions in code distributed with [22], with the binning removed to improve performance for the natural scenes. Fig. 4(e) shows that applying BM3D after variance stabilization using the Anscombe transformation [23], [38] gives a result very similar to Poisson NLSPCA.

To quantify the performance of a reflectivity estimator $\hat{\boldsymbol{\alpha}}$ using reference scene reflectivity $\boldsymbol{\alpha}$, we compute the mean-squared error (MSE) in dB:

$$\text{MSE}(\boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}) = 10 \log_{10} \left(\frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_{i,j} - \hat{\alpha}_{i,j})^2 \right).$$

The values listed in Fig. 4(c) and (f) show that the MSE of our reflectivity estimation method is lower than that of pixel-wise estimation by 17.4 dB. Also, referring to Fig. 4(d) and (e), our computational imager gives MSE reductions of 1.7 dB and 1.1 dB over the state-of-the-art denoising methods of Poisson NLSPCA and BM3D with variance stabilization, respectively. Here and in all subsequent reflectivity experiments, the regularization parameter β_α of our proposed reflectivity imager is chosen to minimize MSE.

B. *Depth Resolution Test*

Depth resolution was evaluated with a test target comprising 5 cm × 5 cm squares of varying thickness mounted on a flat board, as shown by the red-labeled squares in Fig. 4(g). The smallest resolvable height (thickness) above the reference level is an indicator of achievable depth resolution. Fig. 4(l) shows that our method (Equation (16) with the total variation semi-norm for $\text{pen}_z(\cdot)$) achieves 4 mm depth resolution, which is comparable to that of the reference (ground-truth) image (Fig. 4(h)) formed from 100 detections per pixel, and far superior to the log-matched filter image (Equation (13)). Fig. 4(j) and (k), which are the images obtained by applying the median filter and BM3D respectively on the log-matched filter estimate, show that the state-of-the-art denoising methods fail to give significant reduction in RMSE due to the high-variance noise component of the log-matched filter output.

We quantified the performance of a depth estimator $\hat{\mathbf{z}}$ of a true scene depth $\mathbf{z}$ using root mean-square error (RMSE):

$$\text{RMSE}(\mathbf{z}, \hat{\mathbf{z}}) = \sqrt{\frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (z_{i,j} - \hat{z}_{i,j})^2}.$$

At the background level in our experiment, the log-matched filter estimates have an RMSE of at least 3 m. Because many pixels are missing photon detection-time observations, before applying the denoising algorithms, we first fill in every pixel

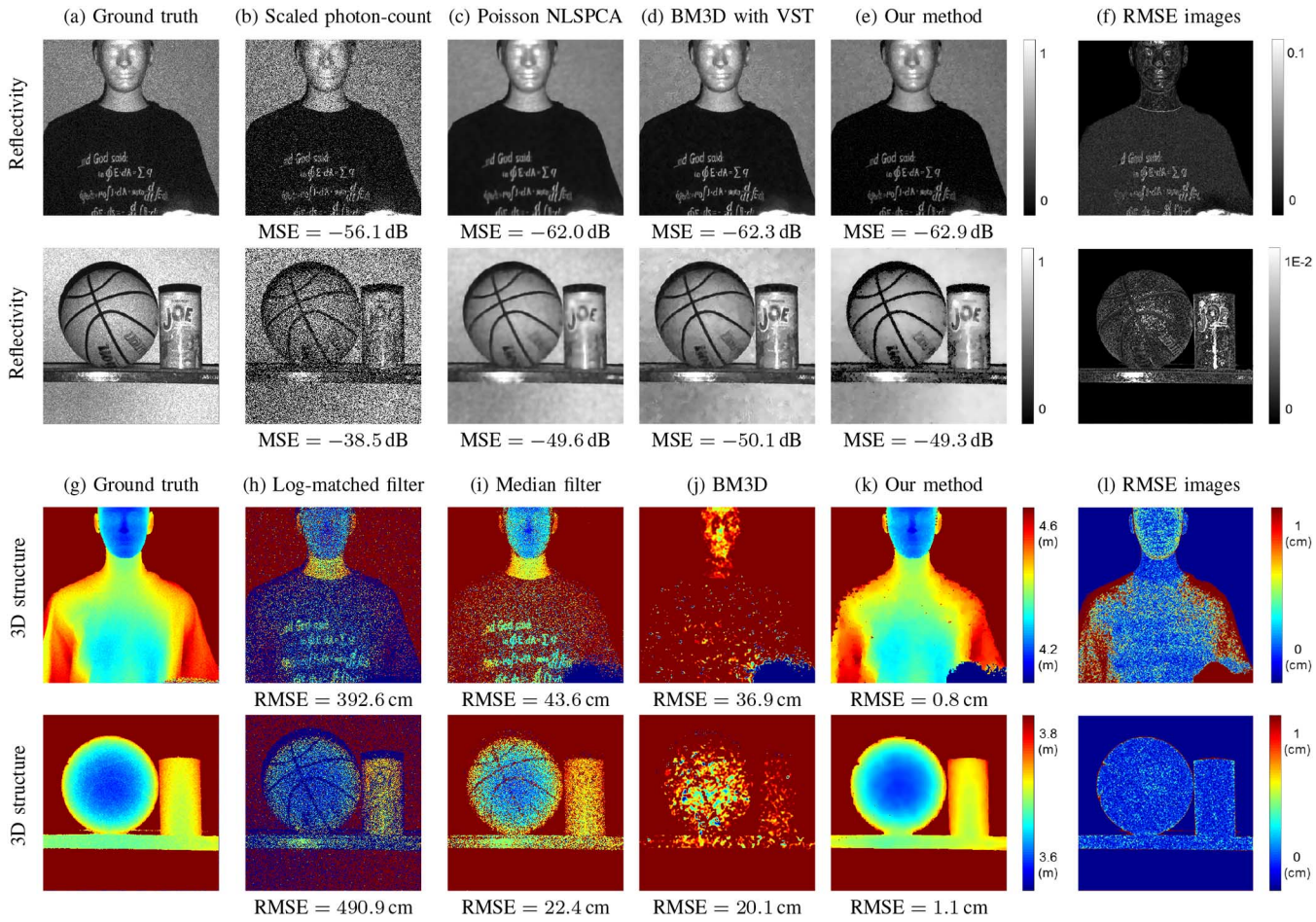


Fig. 5. Experimental results for reflectivity and 3D imaging of natural scenes. All reflectivity images were scaled such that the relevant features are captured in the interval [0,1]. We compare the reflectivity and depth images from our proposed method with those from pixelwise estimation (see Section IV). Pixelwise RMSEs for the reflectivity and 3D images using our method were generated from 100 trials of the experiments. For the mannequin dataset (top), the mean per-pixel count was 1.21 and 54% of the pixels were missing data. For the basketball-and-can dataset (bottom), the mean per-pixel count was 2.1 and 32% of the pixels were missing data.

with missing data with the average of valid depth values in its 5×5 window. Here and in all subsequent depth experiments, the sigma parameter of BM3D was chosen as $\sqrt{1/12}$ based on the fact that the depth from a background noise detection is uniform in [0,1] after scaling the log-matched filter depth image to map detections times in $[0, T_r]$ to [0,1]. This noise standard deviation models the pixels with low reflectivity, while using the square root of the spatial average of the noise variance would be smaller by about a factor of $1/\sqrt{2}$. Also, the regularization parameter β_z of the proposed depth imager is chosen to minimize RMSE. The depth resolution of our method (4 mm) corresponds to more than a 10-fold depth error reduction, compared to the denoised estimates.

C. Natural Scenes

Reflectivity and depth images of two natural scenes—a life-size mannequin, and a basketball next to a can—are shown in Fig. 5. Images that play the role of ground truth were obtained using pixelwise estimation from 200 detections at each pixel and appear in Fig. 5(a) and (g). The mannequin dataset for evaluation of our method was generated using acquisition time $T_a = 100$ µs. This dataset had 1.21 detections per pixel averaged over the entire scene, with 54% of the pixels having no detections. The formation of the reference (ground-truth) mannequin images required an acquisition time of $T_a = 20$ ms. The basketball-and-can dataset for evaluation of our method also had $T_a = 100$ µs, but its mean number of detections per pixel was 2.1, with 32% of the pixels having no detections. The formation of reference (ground-truth) basketball-and-can images required an acquisition time of $T_a = 10$ ms.

Fig. 5(b) and (h) shows that the pixelwise estimation approaches (normalized photon-count (10) for reflectivity and log-matched filter (13) for depth) give estimates with high errors due to background-count shot noise at low light levels. Because of the low light levels, many pixels were missing photon arrivals. For the computation of RMSE of log-matched filter depth, we make a random guess of depth from the interval $[0, cT_r/2)$ at every pixel with missing data.

Denoising the scaled photon-count reflectivity estimates of mannequin and basketball-and-can scenes using Poisson NLSPCA and BM3D with variance stabilization improves the image qualities enough that the resulting images are perceptually similar to the ground-truth images, and MSE reduction is

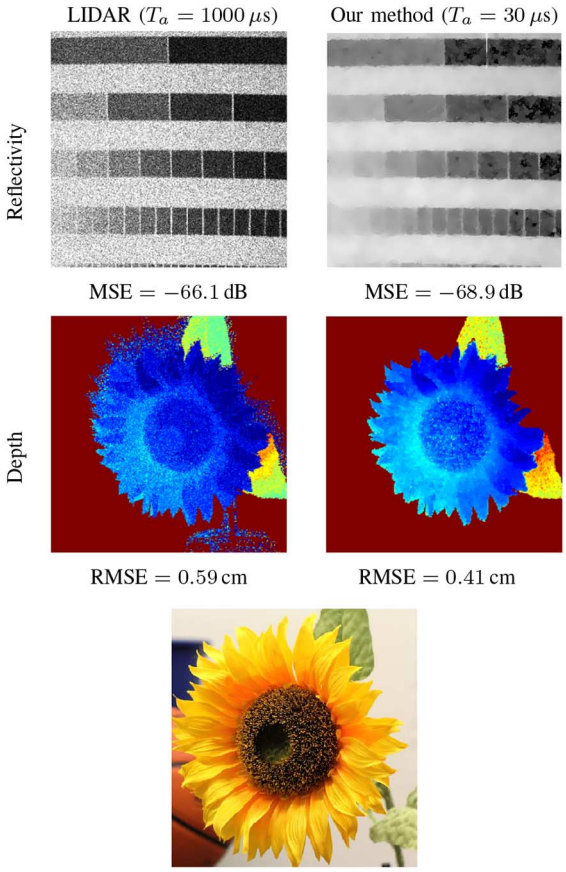


Fig. 6. Comparison between our framework and conventional LIDAR technology.

similar to our method. As in Section VI-B, before denoising the log-matched filter depth image, each pixel with missing data was imputed with the average of valid depth values in its 5×5 window. We observe that denoising the pixelwise depth estimates performs poorly; a large portion of the scene has high background noise and thus many of its pixels have missing data. On the other hand, our framework, which exploits spatial prior information in its noise censoring step, constructs 3D images with small RMSE values of 0.8 cm and 1.1 cm for the mannequin and basketball-and-can scenes, respectively (Fig. 5(k)).

Fig. 6 shows how much photon efficiency we gain over traditional LIDAR systems that use the histogramming approach. The histogramming approach is a pixelwise depth-estimation method that simply searches for the location of the peak in the photon-count histogram of the backreflected pulse. Whereas depth estimation using the log-matched filter is asymptotically ML as $B \to 0^+$, histogramming-based depth estimation method is asymptotically ML as $N \to \infty$. Thus, when T_a is long enough, as is the case in traditional LIDAR, it is effective to use the histogramming-based depth estimation method. Based on values of reflectivity MSE and depth RMSE, we see that our framework can allow more than $30\times$ speed-up in acquisition, while constructing the same high-quality 3D and reflectivity images that a traditional LIDAR system would have formed using long acquisition times.

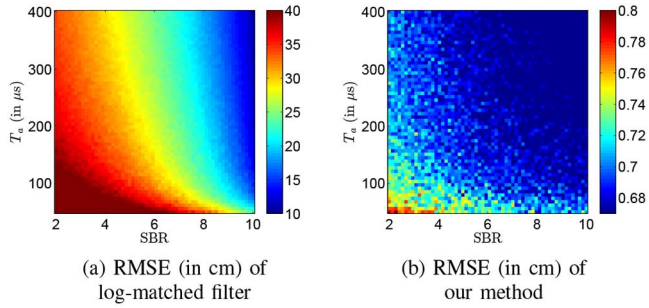


Fig. 7. RMSE simulation results for 3D imaging. Signal-to-background ratio (SBR) given in (17) was varied by simulating background levels on the ground-truth mannequin dataset. Note the differences in the colorbar scales.

D. Repeatability Test

For each natural scene, we processed 100 independent datasets and computed the sample RMSE images that approximate $\sqrt{\mathbb{E}[(\alpha_{i,j} - \hat{\alpha}_{i,j}^{\mathrm{PML}})^2]}$ and $\sqrt{\mathbb{E}[(z_{i,j} - \hat{z}_{i,j}^{\mathrm{PML}})^2]}$ at each pixel (i, j). The pixelwise RMSE images for the mannequin and basketball-and-can scenes are provided in Fig. 5(f) and (l). They corroborate the consistent accuracy and high resolution of our computational reflectivity and 3D imager.

E. Effect of System Parameters

Fig. 7 shows how the performance of traditional pixelwise log-matched filter and our image formation methods are affected by changing the acquisition time T_a and the signal-to-background ratio (SBR), defined to be

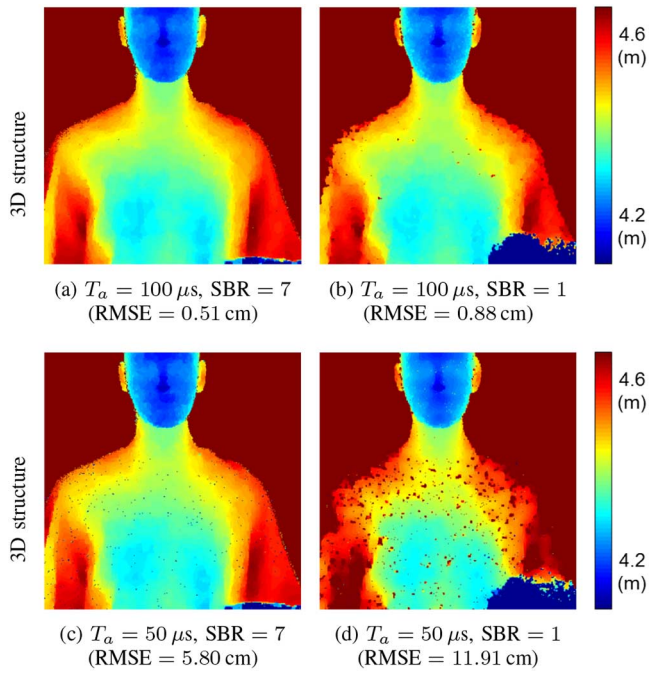
$$\mathrm{SBR} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\eta \alpha_{i,j} S}{B}. \tag{17}$$

To obtain the results reported in Fig. 7, SBR was modified by simulating increases in B through the addition of pseudo-random detections at times uniformly distributed over $[0, T_r]$. RMSE decreases monotonically with increasing T_a and SBR, as one would expect; the RMSE of our method is dramatically lower over all parameter combinations tested.

Fig. 8 provides additional evidence that our 3D recovery method is robust under strong background noise and short acquisition times. Here, SBR was varied by emulating a 7-fold reduction in T_r through time gating; since $B = (\eta b_\lambda + d)T_r$, this emulates a 7-fold increase in SBR.

F. Comparison with First-Photon Imaging

First-photon imaging [3] requires a single detection at each pixel, hence its dwell time on each pixel is a random variable. Our method requires a fixed dwell time on each pixel, hence the number of detections on each pixel is a random variable. So, to compare the performance of first-photon imaging with that of our method, we set the average per-pixel dwell time of the former equal to the fixed per-pixel dwell time of the latter. That comparison, shown in Table I, between the MSEs of their reflectivity images and the RMSEs of their depth images, reveals several interesting characteristics. In particular, when



(a) $T_a = 100\,\mu$s, SBR $= 7$ (RMSE $= 0.51$ cm)
(b) $T_a = 100\,\mu$s, SBR $= 1$ (RMSE $= 0.88$ cm)
(c) $T_a = 50\,\mu$s, SBR $= 7$ (RMSE $= 5.80$ cm)
(d) $T_a = 50\,\mu$s, SBR $= 1$ (RMSE $= 11.91$ cm)

Fig. 8. Effect of dwell time T_a and signal-to-background ratio (SBR) on our 3D recovery method. For acquisition times of 100 µs and 50 µs, we calculated the mean photon count $k_{i,j}$ over all pixels to be 1.4 and 0.6, respectively.

TABLE I
COMPARISON BETWEEN FIRST-PHOTON IMAGING AND OUR FRAMEWORK. NOTE THAT $k_{i,j}$ IS FIXED AND T_a PER PIXEL IS A RANDOM VARIABLE FOR FPI, WHEREAS $k_{i,j}$ IS A RANDOM VARIABLE AND T_a PER PIXEL IS FIXED FOR OUR FRAMEWORK

		FPI	Ours
Mannequin	Mean T_a	244 μs	244 μs
	Mean $k_{i,j}$	1 ppp	2.7 ppp
	Pixels missing data	0%	33%
	MSE of reflectivity	-67 dB	-69 dB
	RMSE of depth	0.4 cm	0.3 cm
Sunflower	Mean T_a	15 μs	15 μs
	Mean $k_{i,j}$	1 ppp	8.7 ppp
	Pixels missing data	0%	18%
	MSE of reflectivity	-53 dB	-54 dB
	RMSE of depth	0.8 cm	0.5 cm
Basketball-and-can	Mean T_a	181 μs	181 μs
	Mean $k_{i,j}$	1 ppp	1.7 ppp
	Pixels missing data	0%	24%
	MSE of reflectivity	-49 dB	-49 dB
	RMSE of depth	1.1 cm	1.1 cm
Reflectivity chart	Mean T_a	120 μs	120 μs
	Mean $k_{i,j}$	1 ppp	1.7 ppp
	Pixels missing data	0%	27%
	MSE of reflectivity	-68 dB	-70 dB
Depth chart	Mean T_a	6.2 μs	6.2 μs
	Mean $k_{i,j}$	1 ppp	1.1 ppp
	Pixels missing data	0%	35%
	RMSE of depth	0.4 cm	0.4 cm

our method's image acquisition time is matched to that of first-photon imaging, a substantial fraction of its pixels have missing data (no detections). Nevertheless, our method successfully deals with this problem and yields performance similar to, or slightly better than, that of first-photon imaging for the five different scenes we have measured.

G. Limitations

Our method of estimating reflectivity becomes less effective as the probability of a photon detection from one pulse grows (i.e., $P_0(\alpha_{i,j})$ in (4) shrinks); this is simply because the dynamic range of $\{k_{i,j}\}_{i,j=1}^n$ shrinks. If the probability of a click is high because of a strong backreflected signal, accurate imaging can be achieved with traditional optical sensing combined with Poisson denoising algorithms [22], [39], [40]. If it is due to high background, performance may be improved by using a narrower spectral filter.

Our method's depth image incurs its highest error near the edges of scene objects. The surface normals at these locations are nearly perpendicular to the line of sight, which dramatically reduces SNR. Consequently, these regions have fewer detections, with more of them being background counts, than do the rest of the pixels. Although our method censors depth anomalies near edges, it estimates the missing depths using spatial correlations, leading to loss of subtle depth details. Also, a detected photon may have originated from a multiple reflection, causing estimation inaccuracy. However, for quasi-Lambertian scenes, diffuse scattering causes multiple reflections to be considerably weaker than the direct reflection. Combined with Poisson statistics, this implies an exponentially diminishing probability of photon detection originating from multiple reflections.

Transverse spatial resolution for both reflectivity and depth is also limited by the accuracy of the raster scanning. Careful examination of the data revealed possible vertical offsets of odd- and even-indexed columns on the order of 2 pixels. With appropriate calibration targets, one may measure the accuracy of the raster scanning and devise appropriate mitigation methods. Methods for mitigation of jitter in sampling may also be appropriate. Here, we believe the effect of this inaccuracy would be similar across the various approximation methods. Jitter in raster scanning is obviated by using a detector array.

Recall from Section II that reflectivity image $\boldsymbol{\alpha} \in \mathbb{R}_+^{n\times n}$ is defined to include the effects of radial fall-off, view angle, and material properties. With the simultaneous estimation of $\boldsymbol{\alpha}$ and depth image $\mathbf{z}$, one may compensate for radial fall-off to obtain estimates of absolute scene patch scattering properties.

VII. CONCLUSIONS AND FUTURE WORK

We have introduced a method for 3D and reflectivity imaging similar to FPI [3]. Both methods use physics-based probabilistic models for the detection of single photons to treat an individual laser pulse as a Bernoulli trial. In FPI, these Bernoulli trials are repeated until the first detection, yielding a geometric random variable; in the present method, a fixed number of Bernoulli trials yields a binomial random variable. This distinction necessitates distinct algorithms and leads to different imaging results.

As a technological matter, using a fixed, deterministic number of laser pulses is greatly advantageous. Our new method's fixed dwell time makes it easier to implement with raster scanning and also compatible with detector arrays. The performance we demonstrate using on the order of 1 detected photon per pixel, averaged over the scene, motivates the development of

accurate and low-power SPAD array-based 3D and reflectivity imagers. Current commercial CMOS-based depth imagers, for example Kinect and TOF cameras, have significantly impacted research in 3D imaging. These sensors offer high depth resolution, but their use is limited due to poor spatial resolution and high power consumption. Our approach offers a potential route to solving these problems.

More generally, our framework can be used in a variety of low-light imaging applications using photon-counting detectors, such as spatially-resolved fluorescence lifetime imaging (FLIM) [41] and high-resolution LIDAR [11]. Our method naturally extends to imaging at a variety of wavelengths, making it suitable for practical implementations. Furthermore, future advances in optoelectronic methods can improve the accuracy of our 3D and reflectivity imager. In particular, it can benefit from improved background suppression techniques [5] and range-gating methods [19].

Many methods in the image processing literature are more sophisticated and powerful than our method in their exploitation of spatial structure. Poisson NLSPCA [22] and BM3D [23], [24] are among these methods and have been used here only for post-processing after pixelwise estimation. More thorough integration of state-of-the-art spatial models with the physical models presented here would presumably provide improved performance.

APPENDIX

This appendix provides performance analyses for pixelwise estimation. The Cramér-Rao lower bound (CRLB) sets the limit on the mean-square error (MSE) of an unbiased estimator of a parameter. Let x be a scalar continuous parameter in the PDF $f_Y(y;x)$ of random variable Y. The CRLB for an unbiased estimator $\hat{x}$ of the parameter x based on observation of Y is the inverse of the Fisher information $J(x)$ [31]:

$$\begin{aligned}\mathbb{E}[(x-\hat{x})^2] &\geq \text{CRLB}(x) = J^{-1}(x) \\ &= \left\{\mathbb{E}\left[\frac{d^2}{d^2x}\left(-\log f_Y(y;x)\right)\right]\right\}^{-1}. \end{aligned} \tag{18}$$

An unbiased estimator $\hat{x}$ is efficient if $\mathbb{E}[(x-\hat{x})^2] = \text{CRLB}(x)$.

A. Mean-Square Error of Reflectivity Estimation

With some algebra, the CRLB for estimating the reflectivity $\alpha_{i,j}$ at pixel (i,j) can be shown using (5) to be

$$\begin{aligned}\text{CRLB}\left(\alpha_{i,j}\right) &= \left\{\mathbb{E}\left[\frac{d^2}{d^2\alpha_{i,j}}\left(-\log \Pr[K_{i,j}=k;\alpha_{i,j}]\right)\right]\right\}^{-1} \\ &= \left\{\mathbb{E}\left[\frac{k\eta^2S^2\exp\left[\eta\alpha_{i,j}S+B\right]}{\left(\exp\left[\eta\alpha_{i,j}S+B\right]-1\right)^2}\right]\right\}^{-1} \\ &= \frac{\exp\left[\eta\alpha_{i,j}S+B\right]-1}{N\eta^2S^2} \\ &\approx \frac{\eta\alpha_{i,j}S+B}{N\eta^2S^2}, \end{aligned} \tag{19}$$

where the approximation makes use of the low-flux condition $\eta\alpha_{i,j}S+B \ll 1$ and $\exp(x) \approx 1+x$ for small x. As could easily be expected, increasing the number of pulse repetitions, N, collects more photons and hence decreases the CRLB.

Note, however, that we cannot directly use the CRLB result to lower bound the mean-square error of the unconstrained ML reflectivity estimate $\hat{\alpha}_{i,j}^{\text{ML}}$ given by

$$\hat{\alpha}_{i,j}^{\text{ML}} = \frac{1}{\eta S}\left[\log\left(\frac{N}{N-k_{i,j}}\right) - B\right].$$

This is because the ML estimate is biased, ($\mathbb{E}[\hat{\alpha}_{i,j}^{\text{ML}}] \neq \alpha_{i,j}$):

$$\begin{aligned}\mathbb{E}\left[\hat{\alpha}_{i,j}^{\text{ML}}\right] &= \mathbb{E}\left[\frac{1}{\eta S}\log\left(\frac{N}{N-k_{i,j}}\right) - \frac{B}{\eta S}\right] \\ &= \frac{1}{\eta S}\log N - \frac{1}{\eta S}\mathbb{E}\left[\log\left(N-K_{i,j}\right)\right] - \frac{B}{\eta S} \\ &> \frac{1}{\eta S}\log N - \frac{1}{\eta S}\log\left(N-\mathbb{E}[K_{i,j}]\right) - \frac{B}{\eta S} \\ &= \alpha_{i,j}, \end{aligned}$$

where the strict inequality comes from Jensen's inequality and the fact that the logarithm function is strictly concave.

When $N \to \infty$ and $\eta\alpha_{i,j}S+B \to 0^+$ with $N[1-\exp(\eta\alpha_{i,j}S+B)]$ equal to a constant $C(\alpha_{i,j})$, the ML reflectivity estimate is

$$\hat{\alpha}_{i,j}^{\text{ML}} = \frac{k}{N\eta S} - \frac{B}{\eta S}. \tag{20}$$

In this case, the CRLB equals the MSE of the ML reflectivity estimate,

$$\text{CRLB}(\alpha_{i,j}) = \mathbb{E}\left[\left(\alpha_{i,j}-\hat{\alpha}_{i,j}^{\text{ML}}\right)^2\right] = \frac{1}{N}\left(\frac{\alpha_{i,j}}{\eta S} + \frac{B}{\eta^2S^2}\right).$$

We see that the CRLB expression from the Poisson likelihood is equal to the first-order Taylor expansion of the CRLB expression of the exact binomial likelihood given by (19).

Knowing that the ML solution in the limiting Poisson distribution is unbiased and efficient, we conclude that the maximum likelihood reflectivity estimate $\hat{\alpha}_{i,j}^{\text{ML}}$ is efficient asymptotically as $(\eta\alpha_{i,j}S+B) \to 0^+$ and $N \to \infty$.

B. Mean-Square Error of 3D Estimation

We again assume that $\eta\alpha_{i,j}S+B \to 0^+$ and $N \to \infty$ such that $N[1-\exp(\eta\alpha_{i,j}S+B)]$ is a constant $C(\alpha_{i,j})$. The CRLB for estimating the depth $z_{i,j}$ is then found to be

$$\begin{aligned}\text{CRLB}(z_{i,j}) &= \left\{\mathbb{E}\left[\frac{d^2}{d^2z_{i,j}}\left(-\log f_{T_{i,j}}(\{t_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}};z_{i,j})\right)\right]\right\}^{-1} \\ &= \left\{\mathbb{E}\left[-\sum_{\ell=1}^{k_{i,j}}\frac{d^2}{d^2z_{i,j}}\log f_{T_{i,j}}(t_{i,j}^{(\ell)};z_{i,j})\right]\right\}^{-1} \\ &= \frac{1}{C(\alpha_{i,j})}\left(\int_0^{T_r}\frac{\dot{p}(t;z_{i,j})^2}{p(t;z_{i,j})}\,dt\right)^{-1}, \end{aligned} \tag{21}$$

where

$$p(t; z_{i,j}) = \frac{\lambda_{i,j}(t)}{\int_0^{T_r} \lambda_{i,j}(\tau)\, d\tau}$$

with $\lambda_{i,j}(t)$ the single-pulse rate from (1) and $\dot{p}(t; z_{i,j})$ its derivative with respect to time.

We can exactly compute the MSE expression for certain pulse waveforms. For example, if the illumination waveform is a Gaussian pulse $s(t) \propto \exp[t^2/2T_p^2]$, then using the unconstrained log-matched filter expression, we get

$$\hat{z}_{i,j}^{\mathrm{ML}} = \arg\max_{z_{i,j}} \sum_{\ell=1}^{k_{i,j}} \log\Big[s(t_{i,j}^{(\ell)} - 2z_{i,j}/c)\Big] = \frac{c}{2}\left(\frac{\sum_{\ell=1}^{k_{i,j}} t_{i,j}^{(\ell)}}{k_{i,j}}\right),$$

given $k_{i,j} \geq 1$. If $k_{i,j} = 0$, then a standard pixelwise data imputation is done by making a uniformly random guess over the interval $[0, z_{\max})$. Assuming $B = 0$, the MSE expression can be written as

$$\begin{aligned}
&\mathbb{E}[(z_{i,j} - \hat{z}_{i,j}^{\mathrm{ML}})^2] = \mathbb{E}_{K_{i,j}}\left[\mathbb{E}[(z_{i,j} - \hat{z}_{i,j}^{\mathrm{ML}})^2 \mid K_{i,j}]\right] \\
&= \sum_{k=0}^{\infty} \frac{C^k(\alpha_{i,j}) e^{-C(\alpha_{i,j})}}{k!} \mathbb{E}[(z_{i,j} - \hat{z}_{i,j}^{\mathrm{ML}})^2 \mid K_{i,j} = k] \\
&= e^{-C(\alpha_{i,j})} \left[\left(\frac{cT_r}{2}\right)^2 + \left(z_{i,j} - \frac{cT_r}{4}\right)^2 \right. \\
&\qquad \left. + \sum_{k=1}^{\infty} \frac{C^k(\alpha_{i,j})}{k!} \frac{1}{k} \left(\frac{cT_p}{2}\right)^2\right] \\
&= e^{-C(\alpha_{i,j})} \left(\underbrace{\left(\frac{cT_r}{2}\right)^2 + \left(z_{i,j} - \frac{cT_r}{4}\right)^2}_{\text{random guess error}} \right. \\
&\qquad \left. + \underbrace{\left(\frac{cT_p}{2}\right)^2 \int_0^{C(\alpha_{i,j})} \frac{\exp[\tau] - 1}{\tau}\, d\tau}_{\text{pulse-width error}} \right). \qquad (22)
\end{aligned}$$

As $C(\alpha_{i,j}) \to \infty$, the pulse-width error term in MSE dominates and $\hat{z}_{i,j}^{\mathrm{ML}}$ becomes an efficient estimator.

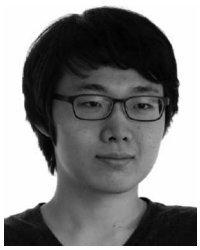
ACKNOWLEDGMENT

The authors would like to thank the Associate Editor A. Foi for many constructive comments and extraordinary attention to detail.

REFERENCES

[1] B. Schwarz, "LIDAR: Mapping the world in 3D," *Nat. Photonics*, vol. 4, no. 7, pp. 429–430, Jul. 2010.
[2] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, "Computational 3D and reflectivity imaging with high photon efficiency," in *Proc. IEEE Int. Conf. Image Process.*, Paris, France, Oct. 2014, pp. 46–50.
[3] A. Kirmani *et al.*, "First-photon imaging," *Science*, vol. 343, no. 6166, pp. 58–61, Jan. 3, 2014.
[4] Massachusetts Institute of Technology, *First-Photon Imaging Project*, 2013 [Online]. Available: http://www.rle.mit.edu/first-photon-imaging/
[5] A. McCarthy, R. J. Collins, N. J. Krichel, V. Fernández, A. M. Wallace, and G. S. Buller, "Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting," *Appl. Opt.*, vol. 48, no. 32, pp. 6241–6251, Nov. 2009.
[6] S. B. Gokturk, H. Yalcin, and C. Bamji, "A time-of-flight depth sensor—System description, issues and solutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2004, p. 35.
[7] S. Lee, O. Choi, and R. Horaud, *Time-of-Flight Cameras: Principles, Methods and Applications*. New York, NY, USA: Springer, 2013.
[8] A. V. Jelalian, "Laser radar systems," in *Proc. Electron. Aerosp. Syst. Conf. (EASCON'80)*, 1980, vol. 1, pp. 546–554.
[9] Z. Zhang, "Microsoft Kinect sensor and its effect," *IEEE Multimedia*, vol. 19, no. 2, pp. 4–10, Apr. 2012.
[10] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.
[11] B. F. Aull *et al.*, "Geiger-mode avalanche photodiodes for three-dimensional imaging," *Lincoln Lab. J.*, vol. 13, no. 2, pp. 335–349, 2002.
[12] A. Kirmani, D. Venkatraman, A. Colaço, F. N. C. Wong, and V. K. Goyal, "High photon efficiency computational range imaging using spatio-temporal statistical regularization," in *Proc. CLEO*, San Jose, CA, USA, Jun. 2013, paper QF1B.2. [Online]. Available: https://www.osapublishing.org/abstract.cfm?uri=CLEO_QELS-2013-QF1B.2
[13] A. Kirmani, A. Colaço, D. Shin, and V. K. Goyal, "Spatio-temporal regularization for range imaging with high photon efficiency," in *Proc. SPIE Wavelets Sparsity XV*, San Diego, CA, USA, Aug. 2013, pp. 88 581F–88 581F.
[14] D. Shin, A. Kirmani, A. Colaço, and V. K. Goyal, "Parametric Poisson process imaging," in *Proc. IEEE Global Conf. Signal Inform. Process.*, Austin, TX, USA, Dec. 2013, pp. 1053–1056.
[15] 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.
[16] A. Colaço, A. Kirmani, G. A. Howland, J. C. Howell, and V. K. Goyal, "Compressive depth map acquisition using a single photon-counting detector: Parametric signal processing meets sparsity," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Providence, RI, USA, Jun. 2012, pp. 96–102.
[17] G. A. Howland, D. J. Lum, M. R. Ware, and J. C. Howell, "Photon counting compressive depth mapping," *Opt. Exp.*, vol. 21, no. 20, pp. 23822–23837, Oct. 2013.
[18] A. Kirmani, A. Colaço, F. N. C. Wong, and V. K. Goyal, "Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor," *Opt. Exp.*, vol. 19, no. 22, pp. 21485–21507, Oct. 2011.
[19] J. Busck and H. Heiselberg, "Gated viewing and high-accuracy three-dimensional laser radar," *Appl. Opt.*, vol. 43, no. 24, pp. 4705–4710, 2004.
[20] G. N. Gol'tsman *et al.*, "Picosecond superconducting single-photon optical detector," *Appl. Phys. Lett.*, vol. 79, no. 6, pp. 705–707, 2001.
[21] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*. Belmont, MA, USA: Athena Scientific, 2002.
[22] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, "Poisson noise reduction with non-local PCA," *J. Math. Image Vis.*, vol. 48, no. 2, pp. 279–294, 2014.
[23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[24] A. Foi, "Noise estimation and removal in MR imaging: The variance-stabilization approach," in *Proc. IEEE Int. Symp. Biomed. Image*, Chicago, IL, USA, Apr. 2011, pp. 1809–1814.
[25] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 839–846.
[26] B. I. Erkmen and B. Moision, "Maximum likelihood time-of-arrival estimation of optical pulses via photon-counting photodetectors," in *Proc. IEEE Int. Symp. Inform. Theory*, 2009, pp. 1909–1913.
[27] D. L. Snyder, *Random Point Processes*. Hoboken, NJ, USA: Wiley, 1975.
[28] D. Shin, "Computational 3D and reflectivity imaging with high photon efficiency," M.S. thesis, Massachusetts Inst. Technol., Cambridge, MA, USA, Jun. 2014.
[29] Y. Chen, J. D. Müller, P. T. So, and E. Gratton, "The photon counting histogram in fluorescence fluctuation spectroscopy," *Biophys. J.*, vol. 77, no. 1, pp. 553–567, Jul. 1999.
[30] F. Villa *et al.*, "SPAD smart pixel for time-of-flight and time-correlated single-photon counting measurements," *IEEE Photonics J.*, vol. 4, no. 3, pp. 795–804, Jun. 2012.

[31] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory. Englewood Cliffs*, NJ, USA: Prentice-Hall, 1998.
[32] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is SPIRAL-TAP: Sparse Poisson intensity reconstruction algorithms—Theory and practice," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1084–1096, Mar. 2012.
[33] E. Abreu, M. Lightstone, S. K. Mitra, and K. Arakawa, "A new efficient approach for the removal of impulse noise from highly corrupted images," *IEEE Trans. Image Process.*, vol. 5, no. 6, pp. 1012–1025, Jun. 1996.
[34] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.
[35] S. Osher, A. Solé, and L. Vese, "Image decomposition and restoration using total variation minimization and the H^{-1} norm," *Multiscale Model. Simul.*, vol. 1, no. 3, pp. 349–370, 2003.
[36] D. Venkatraman (2014). *GitHub Repository for Photon-Efficient Imaging* [Online]. Available: https://github.com/photon-efficient-imaging/sample-data/
[37] A. Kirmani *et al.* (2014). *First-Photon Imaging: Materials/Methods, Supporting Text, Tables, Figures, and/or References* [Online]. Available: http://www.sciencemag.org/content/343/6166/58/suppl/DC1
[38] M. Makitalo and A. Foi, "A closed-form approximation of the exact unbiased inverse of the Anscombe variance-stabilizing transformation," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2697–2698, Sep. 2011.
[39] R. Willett and R. D. Nowak, "Platelets: A multiscale approach for recovering edges and surfaces in photonlimited medical imaging," *IEEE Trans. Med. Imag.*, vol. 22, no. 3, pp. 332–350, Mar. 2003.
[40] M. Mäkitalo and A. Foi, "Optimal inversion of the Anscombe transformation in low-count Poisson image denoising," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 99–109, Jan. 2011.
[41] D. V. O'Connor and D. Phillips, *Time-Correlated Single Photon Counting*. New York, NY, USA: Academic, 1984.



Dongeek Shin (S'13) received the B.S. degrees in electrical engineering and mathematics, both with highest distinction, from the University of Illinois, Urbana-Champaign, in 2012. He received the S.M. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2014. He is currently pursuing the Ph.D. degree in electrical engineering and computer science at MIT. His research interests include statistical inference, sparse representations, and computational imaging. He is a Samsung Fellow. He was also a Finalist in the Qualcomm Innovation Fellowship (QInF) Program in 2013, and was the recipient of the 2014 IEEE International Conference on Image Processing Best Paper Award.

Ahmed Kirmani (S'08) received the Integrated M.Tech. degree in mathematics and computing from the Indian Institute of Technology (IIT) Delhi, New Delhi, India, in 2008, and the S.M. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2010 and 2015, respectively. As a graduate student, he was a Microsoft Research Fellow and a Qualcomm Innovation Fellow. His doctoral research focused on creating novel imaging modalities using computational active optical imaging. He received the Best Paper Award at IEEE ICIP 2014 and the David Marr Prize Honorable Mention at IEEE ICCV 2009. He was also a 2013 Lemelson-MIT Student Prize Finalist.

Vivek K Goyal (S'92–M'98–SM'03–F'14) received the B.S. degree in mathematics and the B.S.E. degree in electrical engineering from the University of Iowa, where he received the John Briggs Memorial Award for the top undergraduate across all colleges. He received the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, CA, USA, where he received the Eliahu Jury Award for outstanding achievement in systems, communications, control, or signal processing. He was a Member of Technical Staff with the Mathematics of Communications Research Department, Bell Laboratories, Lucent Technologies, from 1998 to 2001, and a Senior Research Engineer with Digital Fountain, Inc., from 2001 to 2003. He was with the Massachusetts Institute of Technology from 2004 to 2013, where he was the Esther and Harold E. Edgerton Associate Professor of Electrical Engineering. He is currently an Associate Professor of Electrical and Computer Engineering (on leave) with Boston University. He was an Adviser to 3dim Tech, Inc. (winner of the 2013 MIT $100 K Entrepreneurship Competition Launch Contest Grand Prize), and is now with Google. He is the coauthor of *Foundations of Signal Processing* (Cambridge University Press, 2014). His research interests include computational imaging, human perception, decision making, sampling, quantization, and source coding theory. He is a member of Phi Beta Kappa, Tau Beta Pi, Sigma Xi, Eta Kappa Nu, SIAM, and AAAS. He served on the IEEE Signal Processing Society (SPS) Image and Multiple Dimensional Signal Processing Technical Committee (TC), the SPS Image, Video, and Multidimensional Signal Processing TC, and the Steering Committee of the IEEE TRANSACTIONS ON MULTIMEDIA. He currently serves on the Editorial Board of *Foundations and Trends and Signal Processing*, and the Scientific Advisory Board of the Banff International Research Station for Mathematical Innovation and Discovery. He is a Technical Program Co-Chair of the 11th International Conference on Sampling Theory and Applications, and a permanent Conference Co-Chair of the SPIE Wavelets and Sparsity Conference Series. He was the recipient of the 2002 IEEE Signal Processing Society Magazine Award, the 2014 IEEE International Conference on Image Processing Best Paper Award, and an NSF CAREER Award. The work he supervised won Student Best Paper Awards at the IEEE Data Compression Conference in 2006 and 2011 and the IEEE Sensor Array and Multichannel Signal Processing Workshop in 2012 as well as four MIT Thesis Awards.

Jeffrey H. Shapiro (S'67–M'70–SM'84–F'95–LF'12) received the S.B., S.M., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1967, 1968, 1969, and 1970, respectively. As a graduate student, he was a National Science Foundation Fellow, a Teaching Assistant, and a Fannie and John Hertz Foundation Fellow. His doctoral research was a theoretical study of adaptive techniques for improved optical communication through atmospheric turbulence. From 1970 to 1973, he was an Assistant Professor of Electrical Sciences and Applied Physics with Case Western Reserve University, Cleveland, OH, USA. From 1973 to 1985, he was an Associate Professor of Electrical Engineering with MIT, where he became a Professor of Electrical Engineering in 1985. From 1989 to 1999, he was an Associate Head of the Department of Electrical Engineering and Computer Science. From 2001 to 2011, he was the Director of the Research Laboratory of Electronics. In 1999, he became the Julius A. Stratton Professor of Electrical Engineering. He is best known for his work on the generation, detection, and application of squeezed-state light beams, but he has also authored many works in the areas of atmospheric optical communication, coherent laser radar, and quantum information science. His research interests include the application of communication theory to optical systems. He is a Fellow of the Optical Society of America, the American Physical Society, the Institute of Physics, and SPIE (the International Society for Optical Engineering). In 2008, he was a corecipient of the Quantum Electronics Award from the IEEE Photonics Society (formerly known as IEEE Lasers and Electro-Optics Society), and also received the Quantum Communication Award for Theoretical Research from Tamagawa University, Tokyo, Japan. He has been an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and the *Journal of the Optical Society of America*. He was the Principal Organizer of the Sixth International Conference on Quantum Communication, Measurement and Computing, 2002.