Photon-Flooded Single-Photon 3D Cameras

Anant Gupta Atul Ingle Andreas Velten Mohit Gupta
{agupta225,ingle,velten,mgupta37}@wisc.edu
University of Wisconsin-Madison

Abstract

Single-photon avalanche diodes (SPADs) are starting to play a pivotal role in the development of photon-efficient, long-range LiDAR systems. However, due to non-linearities in their image formation model, a high photon flux (e.g., due to strong sunlight) leads to distortion of the incident temporal waveform, and potentially, large depth errors. Operating SPADs in low flux regimes can mitigate these distortions, but, often requires attenuating the signal and thus, results in low signal-to-noise ratio. In this paper, we address the following basic question: what is the optimal photon flux that a SPAD-based LiDAR should be operated in? We derive a closed form expression for the optimal flux, which is quasi-depth-invariant, and depends on the ambient light strength. The optimal flux is lower than what a SPAD typically measures in real world scenarios, but surprisingly, considerably higher than what is conventionally suggested for avoiding distortions. We propose a simple, adaptive approach for achieving the optimal flux by attenuating incident flux based on an estimate of ambient light strength. Using extensive simulations and a hardware prototype, we show that the optimal flux criterion holds for several depth estimators, under a wide range of illumination conditions.

1. Introduction

Single-photon avalanche diodes (SPAD) are increasingly being used in active vision applications such as fluorescence lifetime-imaging microscopy (FLIM) [33], non-line-of-sight (NLOS) imaging [24], and transient imaging [23]. Due to their extreme sensitivity and timing resolution, these sensors can play an enabling role in demanding imaging scenarios, for instance, long-range LiDAR [6] for automotive applications [20], with only limited power budgets [25].

A SPAD-based LiDAR (Fig. 1) typically consists of a laser which sends out periodic light pulses. The SPAD detects the first incident photon in each laser period, after which it enters a *dead time*, during which it cannot detect any further photons. The first photon detections in each period are then used to create a histogram (over several periods) of the time-of-arrival of the photons. If the incident flux level is sufficiently low, the histogram is approximately a scaled version of the received temporal waveform, and thus, can be used to estimate scene depths and reflectivity.

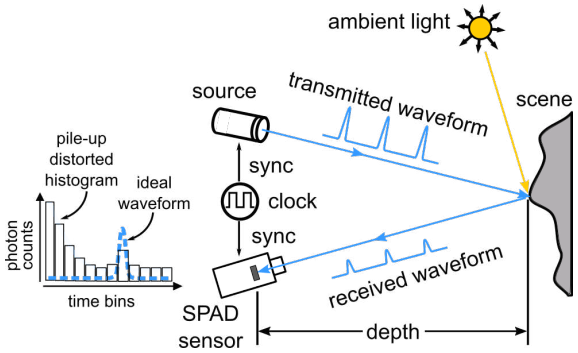

Figure 1. **Pile-up in SPAD-based pulsed LiDAR.** A pulsed LiDAR consists of a light source that illuminates scene points with periodic short pulses. A SPAD sensor records the arrival times of returning photons with respect to the most recent light pulse, and uses those to build a timing histogram. In low ambient light, the histogram is the same shape as the temporal waveform received at the SPAD, and can be used for accurate depth estimation. However, in high ambient light, the histogram is distorted due to pile-up, resulting in potentially large depth errors.

Although SPAD-based LiDARs hold considerable promise due to their single-photon sensitivity and extremely high timing (hence, depth) resolution, the peculiar histogram formation procedure causes severe non-linear distortions due to ambient light [12]. This is because of intriguing characteristics of SPADs under high incident flux: the detection of a photon depends on the time of arrival of previous photons. This leads to non-linearities in the image formation model; the measured histogram gets skewed towards earlier time bins, as illustrated in Figs. 1 and 2[1]. This distortion, also called "pile-up" [12], becomes increasingly severe as the amount of ambient light increases, and can lead to large depth errors. This can severely limit the performance of SPAD-based LiDAR in outdoor conditions, for example, imagine a power-constrained automotive LiDAR operating on a bright sunny day [20].

One way to mitigate these distortions is to attenuate the incident flux sufficiently so that the image formation model becomes approximately linear [26, 13]. However,

[†]This research was supported in part by ONR grants N00014-15-1-2652 and N00014-16-1-2995 and DARPA grant HR0011-16-C-0025.

[1]In contrast, for a conventional, linear-mode LiDAR pixel, the detection of a photon is independent of previous photons (except past saturation). Therefore, ambient light adds a constant value to the entire waveform.

in a LiDAR application, most of the incident flux may be due to ambient light. In this case, lowering the flux (e.g., by reducing aperture size), requires attenuating *both* the ambient and the signal light[2]. While this mitigates distortions, it also leads to signal loss. This fundamental tradeoff between distortion (at high flux) and low signal (at low flux) raises a natural question: Is there an optimal incident flux for SPAD-based active 3D imaging systems?

Optimal incident flux for SPAD-based LIDAR: We address this question by analyzing the non-linear imaging model of SPAD LiDAR. Given a fixed ratio of source-to-ambient light strengths, we derive a closed-form expression for the optimal incident flux. Under certain assumptions, the optimal flux is quasi-invariant to source strength and scene depths, and surprisingly, depends only on the ambient strength and the unambiguous depth range of the system. Furthermore, the optimal flux is lower than that encountered by LiDARs in typical outdoor conditions. This suggests that, somewhat counter-intuitively, reducing the total flux improves performance, even if that means attenuating the signal. On the other hand, the optimal flux is considerably higher than that needed for the image formation to be in the linear regime [2, 15]. As a result, while the optimal flux still results in some degree of distortion, with appropriate computational depth estimators, it achieves high performance across a wide range of imaging scenarios.

Based on this theoretical result, we develop a simple adaptive scheme for SPAD LiDAR where the incident flux is adapted based on an estimate of the ambient light strength. We perform extensive simulation and hardware experiments to demonstrate that the proposed approach achieves up to an order of magnitude higher depth precision as compared to existing *rule-of-thumb* approaches [2, 15] that require lowering flux levels to linear regimes.

Implications: The theoretical results derived in this paper can lead to a better understanding of this novel and exciting sensing technology. Although our analysis is performed for an analytical pixel-wise depth estimator [7], we show that in practice, the improvements in depth estimation are achieved for several reconstruction approaches, including pixel-wise statistical approaches such as MAP, as well as estimators that account for spatial correlations and scene priors (e.g., neural network estimators [17]). These results may motivate the design of practical, low-power LiDAR systems that can work in a wide range of illumination conditions, ranging from dark to extreme sunlight.

2. Related Work

SPAD-based active vision systems: Most SPAD-based LiDAR, FLIM and NLOS imaging systems [5, 16, 34, 29, 17, 3] rely on the incident flux being sufficiently low so that pile-up distortions can be ignored. Recent work [13] has addressed the problem of source light pile-up for SPAD-based LiDAR using a realistic model of the laser pulse shape and statistical priors on scene structure to achieve sub-pulse-width depth precision. Our goal is different—we provide theoretical analysis and design of SPAD LiDAR that can perform robustly even in strong *ambient* light.

Theoretical analysis and computational methods for pile-up correction: Pile-up distortion can be removed in post-processing by computationally inverting the non-linear image formation model [7, 35]. While these approaches can mitigate relatively low amount of pile-up, they have limited success in high flux levels, where a computational approach alone results in strong amplification of noise. Previous work has performed theoretical analysis similar to ours in a range-gating scenario where scene depths are known [10, 36, 9]. In contrast, we derive an optimal flux criterion that minimizes pile-up errors at capture time, is applicable for a broad range of, including extremely high, lighting levels, and does not require prior knowledge of scene depths.

Alternative sensor architectures: Pile-up can be suppressed by modifying the detector hardware, eg. by using multiple SPADs per pixel connected to a single time-correlated single-photon counting (TCSPC) circuit to distribute the high incident flux over multiple SPADs [3]. Multi-SPAD schemes with parallel timing units and multi-photon thresholds can be used to detect correlated signal photons [28] and reject ambient light photons that are temporally randomly distributed. The theoretical criteria derived here can be used in conjunction with these hardware architectures for optimal LiDAR design.

Active 3D imaging in sunlight: Prior work in the structured light and time-of-flight literature proposes various coding and illumination schemes to address the problem of low signal-to-noise ratios (SNR) due to strong ambient light [18, 11, 22, 1]. The present work deals with a different problem of optimal photon detection for SPAD-based pulsed time-of-flight. These previous strategies can potentially be applied in combination with our method to further improve depth estimation performance.

3. Background: SPAD LiDAR Imaging Model

This section provides mathematical background on the image formation model for SPAD-based pulsed LiDAR. Such a system typically consists of a laser source which transmits periodic short pulses of light at a scene point, and a co-located SPAD detector [21, 31, 8] which observes the reflected light, as shown in Fig. 1. We model an ideal laser pulse as a Dirac delta function $\tilde{\delta}(t)$. Let d be the distance of the scene point from the sensor, and $\tilde{\tau} = {}^{2d}/_{c}$ be the round trip time-of-flight for the light pulse. The photon flux incident on the SPAD is given by:

$$\Phi(t) = \tilde{\Phi}_{\text{sig}}\, \tilde{\delta}(t - \tilde{\tau}) + \tilde{\Phi}_{\text{bkg}}, \tag{1}$$

where $\tilde{\Phi}_{\text{sig}}$ is the signal component of the received waveform; it encapsulates the laser source power, distance-

[2] Ambient light can be reduced to a limited extent via spectral filtering.

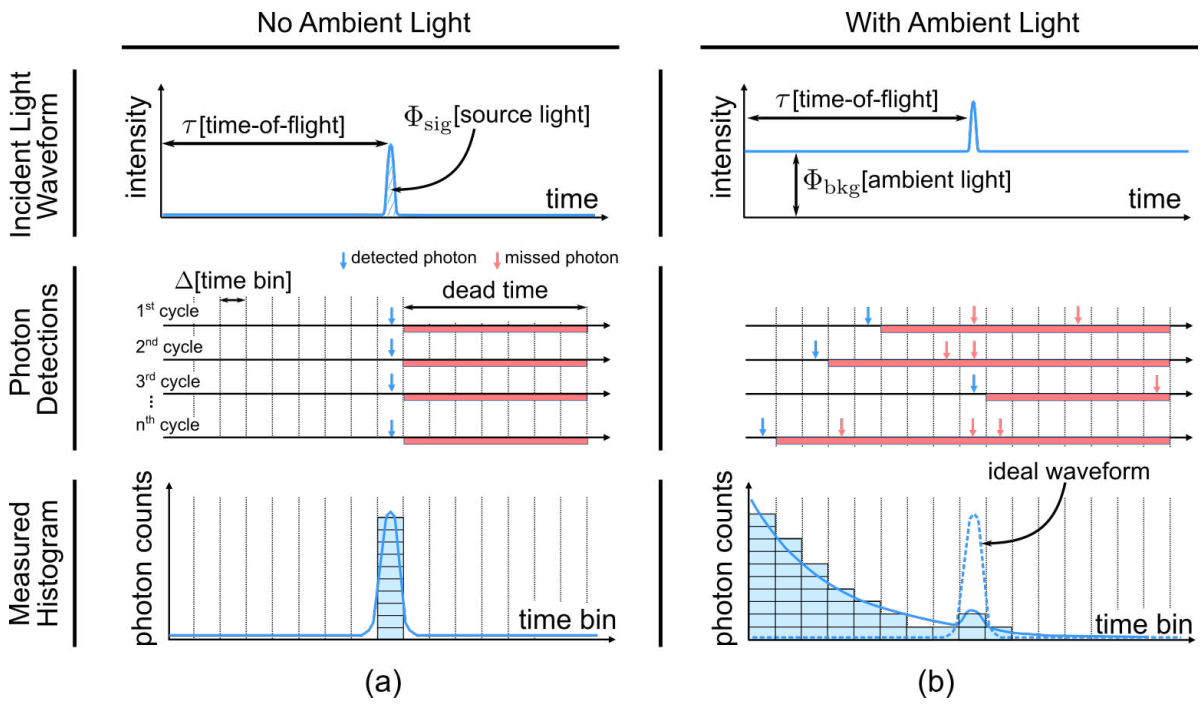

Figure 2. **Effect of ambient light on SPAD LiDAR.** A SPAD-based pulsed LiDAR builds a histogram of the time-of-arrival of the incident photons, over multiple laser pulse cycles. In each cycle, at most one photon is recorded, whose timestamp is used to increment the counts in the corresponding histogram bin. (Left) When there is no ambient light, the histogram is simply a discretized, scaled version of the incident light waveform. (Right) Ambient light photons arriving before the laser pulse skew the shape of the histogram, causing a non-linear distortion, called pile-up. This results in large depth errors, especially as ambient light increases.

squared fall-off, scene brightness and BRDF. $\tilde{\Phi}_{\text{bkg}}$ denotes the background component, assumed to be a constant due to ambient light. Since SPADs have a finite time resolution (few tens of picoseconds), we consider a discretized version of the continuous waveform in Eq. (1), using uniformly spaced time bins of size Δ. Let M_i be the number of photons incident on the SPAD in the i^{th} time bin. Due to arrival statistics of photons, M_i follows a Poisson distribution. The mean of the Poisson distribution, $\mathbb{E}[M_i]$, i.e., the average number r_i of photons incident in i^{th} bin, is given as:

$$r_i = \Phi_{\text{sig}}\, \delta_{i,\tau} + \Phi_{\text{bkg}}. \tag{2}$$

Here, $\delta_{i,j}$ is the Kronecker delta,[3] Φ_{sig} is the mean number of signal photons received per bin, and Φ_{bkg} is the (undesirable) background and dark count photon flux per bin. Let B be the total number of time bins. Then, we define the vector of values $(r_1, r_2, \ldots, r_B)$ as the *ideal* incident waveform.

SPAD histogram formation: SPAD-based LiDAR systems operate on the TCSPC principle [15]. A scene point is illuminated by a periodic train of laser pulses. Each period starting with a laser pulse is referred to as a *cycle*. The SPAD detects only the first incident photon in each cycle, after which it enters a dead time ($\sim$100 ns), during which it cannot detect any further photons. The time of arrival of the first photon is recorded with respect to the start of the most recent cycle. A histogram of first photon arrival times is constructed over many laser cycles, as shown in Fig. 2.

If the histogram consists of B time bins, the laser repetition period is $B\Delta$, corresponding to an unambiguous depth range of $d_{\text{max}} = cB\Delta/2$. Since the SPAD only records the first photon in each cycle, a photon is detected in the i^{th} bin only if at least one photon is incident on the SPAD during the i^{th} bin, and, no photons are incident in the preceding bins. The probability q_i that at least one photon is *incident* during the i^{th} bin can be computed using the Poisson distribution with mean r_i [7]:

$$q_i = \mathsf{P}(M_i \geq 1) = 1 - e^{-r_i}.$$

Thus, the probability p_i of *detecting* a photon in the i^{th} bin, in any laser cycle, is given by [27]:

$$p_i = q_i \prod_{k=1}^{i-1}(1-q_k) = \left(1 - e^{-r_i}\right) e^{-\sum_{k=1}^{i-1} r_k}. \tag{3}$$

Let N be the total number of laser cycles used for forming a histogram and N_i be the number of photons detected in the i^{th} histogram bin. The vector $(N_1, N_2, \ldots, N_{B+1})$ of the histogram counts follows a multinomial distribution:

$$(N_1, N_2, \ldots, N_{B+1}) \sim \mathsf{Mult}(N, (p_1, p_2, \ldots, p_{B+1})), \tag{4}$$

where, for convenience, we have introduced an additional $(B+1)^{\text{st}}$ index in the histogram to record the number of cycles with no detected photons. Note that $p_{B+1} =$

[3] The Kronecker delta is defined as $\delta_{i,j} = 1$ for $i = j$ and 0 otherwise.

$1 - \sum_{i=1}^{B} p_i$ and $N = \sum_{i=1}^{B+1} N_i$. Eq. (4) describes a general probabilistic model for the histogram of photon counts acquired by a SPAD-based pulsed LiDAR.

Fig. 2 (a) shows the histogram formation in the case of negligible ambient light. In this case, all the photon arrival times line up with the location of the peak of the incident waveform. As a result, $r_i = 0$ for all the bins except that corresponding to the laser impulse peak. In this case, the measured histogram vector $(N_1, N_2, \ldots, N_B)$, on average, is simply a scaled version of the incident waveform $(r_1, r_2, \ldots, r_B)$. The time-of-flight can be estimated by locating the bin index with the highest photon counts:

$$\widehat{\tau} = \underset{1 \le i \le B}{\arg\max}\ N_i\,, \tag{5}$$

and the scene depth can be estimated as $\widehat{d} = \frac{c\widehat{\tau}\Delta}{2}$.

For ease of theoretical analysis, we assume the laser pulse is a perfect Dirac-impulse with a duration of a single time bin. We also ignore other SPAD non-idealities such as jitter and afterpulsing. We show in the supplement that the results presented here can potentially be improved by combining our optimal photon flux criterion with recent work [13] that explicitly models the laser pulse shape and SPAD timing jitter.

4. Effect of Ambient Light on SPAD LiDAR

If there is ambient light, the waveform incident on the SPAD can be modeled as an impulse with a constant vertical shift, as shown in the top of Fig. 2 (b). The measured histogram, however, does not reliably reproduce this "DC shift" due to the peculiar histogram formation procedure that only captures the first photon for each laser cycle. When the ambient flux is high, the SPAD detects an ambient photon in the earlier histogram bins with high probability, resulting in a distortion with an exponentially decaying shape. This is illustrated in the bottom of Fig. 2 (b), where the peak due to laser source appears only as a small blip in the exponentially decaying tail of the measured histogram. The problem is exacerbated for scene points that are farther from the imaging system. This distortion, called pile-up, significantly lowers the accuracy of depth estimates because the bin corresponding to the true depth no longer receives the maximum number of photons. In the extreme case, the later histogram bins might receive no photons, making depth reconstruction at those bins impossible.

Computational Pile-up Correction: In theory, it is possible to "undo" the distortion by inverting the exponential nonlinearity of Eq. (3), and finding an estimate of the incident waveform r_i in terms of the measured histogram N_i:

$$\widehat{r}_i = \ln\left(\frac{N - \sum_{k=1}^{i-1} N_k}{N - \sum_{k=1}^{i-1} N_k - N_i}\right). \tag{6}$$

This method is called the Coates's correction [7], and it can be shown to be equivalent to the maximum-likelihood estimate of r_i [27]. See supplementary document for a self-contained proof. The depth can then be estimated as:

$$\widehat{\tau} = \underset{1 \le i \le B}{\arg\max}\, \widehat{r}_i. \tag{7}$$

Although this computational approach removes distortion, the non-linear mapping from measurements N_i to the estimate $\widehat{r}_i$ significantly amplifies measurement noise at later time bins, as shown in Fig. 3.

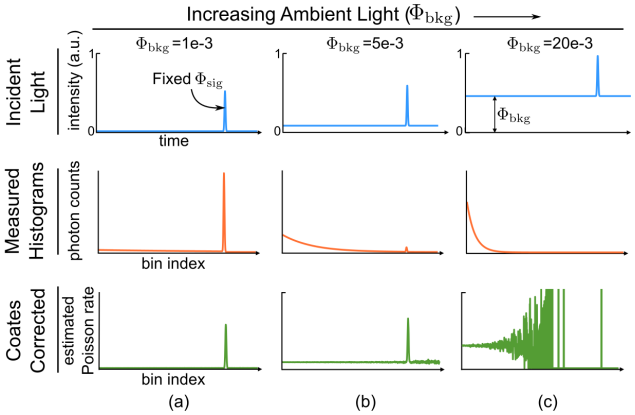

Figure 3. **Efficacy of computational pile-up correction approaches [7].** (a) In low ambient light, there is negligible pile-up. (b) At moderate ambient light levels, pile-up can be observed as a characteristic exponential fall-off in the acquired histogram. The signal pulse location can still be recovered using computational correction (Section 4). (c) In strong ambient light, the later histogram bins receive very few photons, which makes the computationally corrected waveform extremely noisy, making it challenging to reliably locate the laser peak for estimating depth.

Pile-up vs. Low Signal Tradeoff: One way to mitigate pile up is to reduce the total incident photon flux (e.g., by reducing the aperture or SPAD size). Various rules-of-thumb [2, 15] advocate maintaining a low enough photon flux so that only 1-5% of the laser cycles result in a photon being detected by the SPAD. In this case, $r_i \ll 1\,\forall\, i$ and Eq. (3) simplifies to $p_i \approx r_i$. Therefore, the mean photon counts N_i become proportional to the incident waveform r_i, i.e., $\mathbb{E}[N_i] = Np_i \approx Nr_i$. This is called the *linear operation regime* because the measured histogram $(N_i)_{i=1}^{B}$ is, on average, simply a scaled version of the true incident waveform $(r_i)_{i=1}^{B}$. This is similar to the case of no ambient light as discussed above, where depths can be estimated by locating the histogram bin with the highest photon counts.

Although lowering the overall photon flux to operate in the linear regime reduces ambient light and prevents pile-up distortion, unfortunately, it also reduces the source signal considerably. On the other hand, if the incident photon flux is allowed to remain high, the histogram suffers from pile-up, undoing which leads to amplification of noise. This fundamental tradeoff between pile-up distortion and low signal raises a natural question: What is the optimal incident flux level for the problem of depth estimation using SPADs?

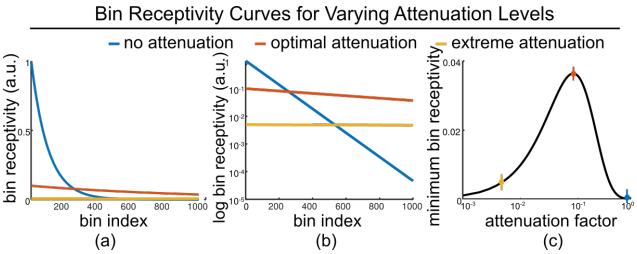

Figure 4. **Bin receptivity curves (BRC) for different attenuation levels.** (a-b) Large (extreme) attenuation results in flat BRC with no pile-up, but low signal level. No attenuation results in a distorted BRC, but higher signal level. The proposed optimal attenuation level achieves a BRC with both low distortion, and high signal. (c) The optimal attenuation factor is given by the maxima location (unique) of the minimum value of BRC.

5. Bin Receptivity and Optimal Flux Criterion

In this section, we formalize the notion of optimal incident photon flux for a SPAD-based LiDAR. We model the *original* incident waveform as a constant ambient light level Φ_{bkg}, with a single source light pulse of height Φ_{sig}. We assume that we can modify the incident waveform only by attenuating it with a scale factor $\Upsilon \leq 1$. This attenuates *both* the ambient Φ_{bkg} and source Φ_{sig} components proportionately. [4] Then, given a Φ_{bkg} and Φ_{sig}, the total photon flux incident on the SPAD is determined by the factor Υ. Therefore, the problem of finding the optimal total incident flux can be posed as determining the optimal attenuation Υ. To aid further analysis, we define the following term.

Definition 1. [Bin Receptivity Coefficient] The *bin receptivity coefficient* C_i of the i^{th} histogram bin is defined as:

$$C_i = \frac{p_i}{r_i} r \,, \tag{8}$$

where p_i is the probability of detecting a photon (Eq. (3)), and r_i is the average number of incident photons (Eq. (2)) in the i^{th} bin. r is the total incident flux $r = \sum_{i=1}^{B} r_i$. The *bin receptivity curve* (BRC) is defined as the plot of the bin receptivity coefficients C_i as a function of the bin index i.

The BRC can be considered an intuitive indicator of the performance of a SPAD LiDAR system, since it captures the pile-up vs. shot noise tradeoff. The first term $\frac{p_i}{r_i}$ quantifies the distortion in the shape of the measured histogram with respect to the ideal incident waveform, while the second term r quantifies the strength of the signal. Figs. 4 (a-b) show the BRCs for high and low incident flux, achieved by using a high and low attenuation Υ, respectively. For small Υ (low flux), the BRC is uniform (negligible pile-up, as $\frac{p_i}{r_i} \approx 1$ is approximately constant across i), but the curve's values are small (low signal). For large Υ (high flux), the curve's values are large on average (large signal), but skewed towards earlier bins (strong pile-up, as $\frac{p_i}{r_i}$ varies considerably from ≈ 1 for earlier bins to $\ll 1$ for later bins). Higher the flux, larger the variation in $\frac{p_i}{r_i}$ over i.

BRC as a function of attenuation factor Υ: Assuming total background flux $B\Phi_{\text{bkg}}$ over the entire laser period to be considerably stronger than the total source flux, i.e., $\Phi_{\text{sig}} \ll B\Phi_{\text{bkg}}$, the flux incident in the i^{th} time bin can be approximated as $r_i \approx r/B$. Then, using Eqs. (8) and (3), the BRC can be expressed as:

$$C_i = B\,(1 - e^{-\frac{r}{B}})\, e^{-(i-1)\frac{r}{B}} . \tag{9}$$

Since total incident flux $r = \Upsilon\,(\Phi_{\text{sig}} + B\,\Phi_{\text{bkg}})$, and we assume $\Phi_{\text{sig}} \ll B\Phi_{\text{bkg}}$, r can be approximated as $r \approx \Upsilon\, B\, \Phi_{\text{bkg}}$. Substituting in Eq. (9), we get an expression for BRC as a function *only* of the attenuation Υ, for a given number of bins B and a background flux Φ_{bkg}:

$$C_i(\Upsilon) = B\,(1 - e^{-\Upsilon\,\Phi_{\text{bkg}}})\, e^{-(i-1)\Upsilon\,\Phi_{\text{bkg}}} . \tag{10}$$

Eq. (10) allows us to navigate the space of BRCs, and hence, the shot noise vs. pile-up tradeoff, by varying a single parameter: the attenuation factor Υ. Based on Eq. (10), we are now ready to define the optimal Υ.

Result 1 (Attenuation and Probability of Depth Error). *Let τ be the true depth bin and $\widehat{\tau}$ the estimate obtained using the Coates's estimator (Eq.(7)). An upper bound on the average probability of depth error $\sum_{\tau=1}^{B} \mathbb{P}(\widehat{\tau} \neq \tau)$ is minimized when the attenuation fraction is given by:*

$$\boxed{\Upsilon^{\text{opt}} = \arg\max_{\Upsilon} \min_{i} C_i(\Upsilon).} \tag{11}$$

See the supplementary technical report for a proof. This result states that, given a signal and background flux, the optimal depth estimation performance is achieved when the minimum bin receptivity coefficient is maximized.

From Eq. (10) we note that for a fixed Υ, the smallest receptivity value is attained at the last bin $i = B$, i.e., $\min_i C_i(\Upsilon) = C_B(\Upsilon)$. Substituting in Eq. (11), we get:

$$\Upsilon^{\text{opt}} = \arg\max_{\Upsilon} C_B(\Upsilon).$$

Using $C_B(\Upsilon)$ from Eq. (10) and solving for Υ, we get:

$$\Upsilon^{\text{opt}} = \frac{1}{\Phi_{\text{bkg}}} \log\left(\frac{B}{B-1}\right).$$

Finally, assuming that $B \gg 1$, we get $\log\left(\frac{B}{B-1}\right) \approx \frac{1}{B}$. Since $B = 2d_{\text{max}}/c\Delta$, where d_{max} is the unambiguous depth range, the final optimality condition can be written as:

$$\underbrace{\boxed{\Upsilon^{\text{opt}} = \frac{c\,\Delta}{2\,d_{\text{max}}\,\Phi_{\text{bkg}}} .}}_{\text{Optimal Flux Attenuation Factor}} \tag{12}$$

Geometric interpretation of the optimality criterion: Result 1 can be intuitively understood in terms of the space of shapes of the BRC. Figs. 4 (a-b) shows the effect of

[4] It is possible to selectively attenuate only the ambient component, to a limited extent, via spectral filtering. We assume that the ambient level Φ_{bkg} is already at the minimum level that is achievable by spectral filtering.

three different attenuation levels on the BRC of a SPAD exposed to high ambient light. When no attenuation is used, the BRC decays rapidly due to strong pile-up. Current approaches [2, 15] that use *extreme attenuation* [5] make the BRC approximately uniform across all histogram bins, but very low on average, resulting in extremely low signal. With optimal attenuation, the curve displays some degree of pile-up, albeit much lower distortion than the case of no attenuation, but considerably higher values, on average, compared to extreme attenuation. Fig. 4 (c) shows that the optimal attenuation factor is given by the unique maxima location of the minimum value of BRC.

Choice of optimality criterion: Ideally, we should minimize the root-mean-squared depth error (RMSE or L^2) in the design of optimal attenuation. However, this leads to an intractable optimization problem. Instead, we choose an upper bound on mean probability of depth error (L^0) as a surrogate metric, which leads to a closed form minimizer. Our simulations and experimental results show that even though Υ^{opt} is derived using a surrogate metric, it also approximately minimizes L^2 error, and provides nearly an order of magnitude improvement in L^2 error.

Estimating Φ_{bkg}: In practice, Φ_{bkg} is unknown and may vary for each scene point due to distance and albedo. We propose a simple adaptive algorithm (see supplement) that first estimates Φ_{bkg} by capturing data over a few initial cycles with the laser source turned off, and then adapts the attenuation at each point by using the estimated Φ_{bkg} in Eq. (11) on a per-pixel basis.

Implications of the optimality criterion: Note that Υ^{opt} is quasi-invariant to scene depths, number of cycles, as well as the signal strength Φ_{sig} (assuming $\Phi_{\text{sig}} \ll B\Phi_{\text{bkg}}$). Depth-invariance is by design—the optimization objective in Result 1 assumes a uniform prior on the true depth. As seen from Eq. (11), this results in an Υ^{opt} that doesn't depend on any prior knowledge of scene depths, and can be easily computed using quantities that are either known (Δ and d_{max}) or can be easily estimated in real-time (Φ_{bkg}). The optimal attenuation fraction can be achieved in practice using a variety of methods including aperture stops, varying the SPAD quantum efficiency, or with ND-filters.

6. Empirical Validation using Simulations

Simulated single-pixel mean depth errors: We performed Monte Carlo simulations to demonstrate the effect of varying attenuation on the mean depth error. We assumed a uniform depth distribution over a range of 1000 time bins.

[5] For example, consider a depth range of 100 m and a bin resolution of $\Delta = 100$ ps. Then, the 1% rule of thumb recommends extreme attenuation so that each bin receives $\approx 1.5 \times 10^{-6}$ photons. In contrast, the proposed optimality condition requires that, on average, one background photon should be incident on the SPAD, per laser cycle. This translates to $\approx 1.5 \times 10^{-4}$ photons per bin, which is orders of magnitude higher than extreme attenuation, and, results in considerably larger signal and SNR.

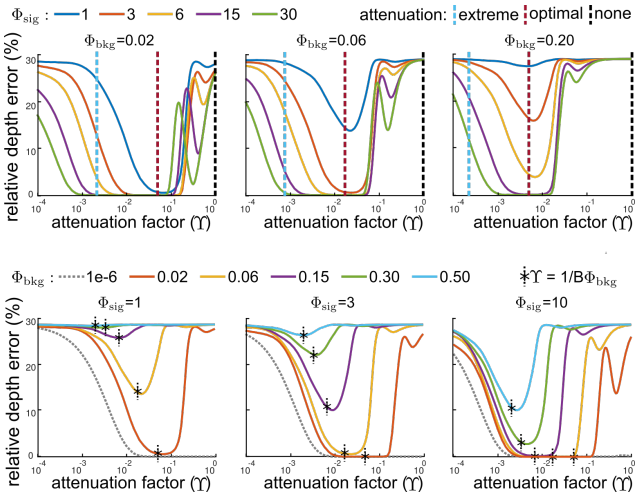

Figure 5. **Simulation based validation.** (Top row) The values of no, extreme, and optimal attenuation are indicated by dotted vertical lines. In each of the three plots, the value of optimal attenuation is approximately invariant to source power level. The optimal attenuation factor depends only on the fixed ambient light level. (Bottom row) For fixed values of source power, the optimal attenuation factor increases as ambient light decreases. The locations of theoretically predicted optimal attenuation (dotted vertical lines) line up with the valleys of the depth error curves.

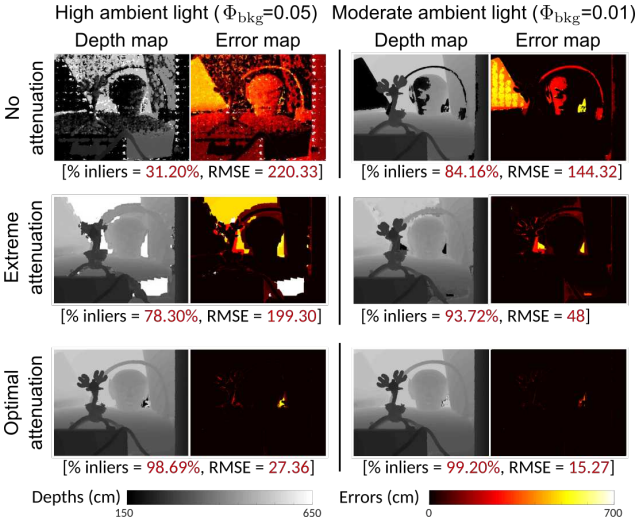

Figure 6. **Neural network based reconstruction for simulations.** Depth and error maps for neural networks-based depth estimation, under different levels of ambient light and attenuation. Extreme attenuation denotes average $\Upsilon B\Phi_{\text{bkg}} = 0.05$. Optimal attenuation denotes $\Upsilon B\Phi_{\text{bkg}} = 1$. % inliers denotes the percentage of pixels with absolute error < 36 cm. $\Phi_{\text{sig}} = 2$ for all cases.

Eq. (6) was used to estimate depths. Fig. 5 shows plots of the relative RMSE as a function of attenuation factor Υ, for a wide range of Φ_{bkg} and Φ_{sig} values.

Each plot in the top row corresponds to a fixed ambient flux Φ_{bkg}. Different lines in a plot correspond to different signal flux levels Φ_{sig}. There are two main observations

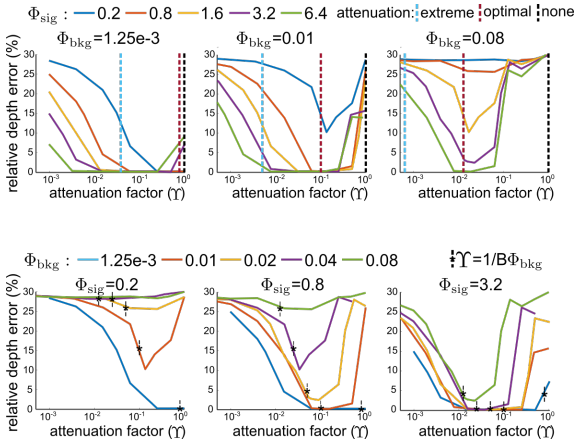

Figure 7. **Validation of optimal attenuation using hardware experiments.** These plots have the same layout as the simulations of Fig. 5. As in simulations, the theoretically predicted locations of the optimal attenuation match the valleys of the depth error curves.

to be made here. First, the optimal attenuation predicted by Eq. (12) (dotted vertical line) agrees with the locations of the minimum depth error valleys in these error plots.[6] Second, the optimal attenuation is quasi-independent of the signal flux Φ_{sig}, as predicted by Eq. (12). Each plot in the second row corresponds to a fixed source flux Φ_{sig}; different lines represent different ambient flux levels. The predicted optimal attenuation align well with the valleys of respective lines, and as expected, are different for different lines.

Improvements in depth estimation performance: As seen from all the plots, the proposed optimal attenuation criterion can achieve up to *1 order of magnitude* improvement in depth estimation error as compared to extreme or no attenuation. Since most valleys are relatively flat, in general, the proposed approach is robust to uncertainties in the estimated background flux, and thus, can achieve high depth precision across a wide range of illumination conditions.

Validation on neural networks-based depth estimation: Although the optimality condition is derived using an analytic pixel-wise depth estimator [7], in practice, it is valid for state-of-the-art deep neural network (DNN) based methods that exploit spatio-temporal correlations in natural scenes. We trained a convolutional DNN [17] using simulated pile-up corrupted histograms, generated using ground truth depth maps from the NYU depth dataset V2 [19], and tested on the Middlebury dataset [32]. For each combination of ambient flux, source flux and attenuation factor, a separate instance of the DNN was trained on corresponding training data, and tested on corresponding test data to ensure a fair comparison across the different attenuation methods.

Fig. 6 shows depth map reconstructions at different levels of ambient light. If no attenuation is used with high ambient light, the acquired data is severely distorted by pile-up, resulting in large depth errors. With extreme attenuation, the DNN is able to smooth out the effects of shot noise, but results in blocky edges. With optimal attenuation, the DNN successfully recovers the depth map with considerably higher accuracy, at all ambient light levels.

[6] As explained in the supplement, the secondary dips in these error plots at high flux levels are an artifact of using the Coates's estimator, and are removed by using more sophisticated estimators such as MAP.

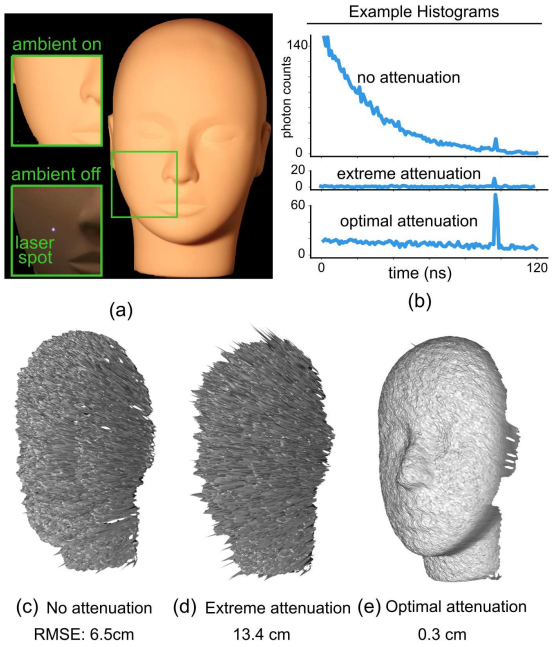

Figure 8. **3D reconstruction of a mannequin face** (a) A mannequin face illuminated by bright ambient light. The laser spot is barely visible. (b) Representative histograms acquired from the laser position shown in (a). With extreme and no attenuation, the peak corresponding to the scene depth is barely identifiable. With optimal attenuation, the peak can be extracted reliably. (c-d) The depth reconstructions using no and extreme attenuation suffer from strong pile-up and shot noise, (e) Optimal attenuation achieves an order of magnitude higher depth precision, even enabling recovery of fine details.

7. Hardware Prototype and Experiments

Our hardware prototype is similar to the schematic shown in Fig. 1. We used a $405\,\text{nm}$ wavelength, pulsed, picosecond laser (PicoQuant LDH P-C-405B) and a co-located fast-gated single-pixel SPAD detector [4] with a $200\,\text{ns}$ dead time. The laser repetition rate was set to 5 MHz corresponding to $d_{\text{max}} = 30\,\text{m}$. Photon timestamps were acquired using a TCSPC module (PicoQuant HydraHarp 400). Due to practical space constraints, various depths covering the full $30\,\text{m}$ of unambiguous depth range in Fig. 7 were emulated using a programmable delayer module (Micro Photon Devices PSD). Similarly, all scenes in Figs. 8, 9 and 10 were provided with a depth offset of $15\,\text{m}$ using the PSD, to mimic long range LiDAR.

Single-pixel Depth Reconstruction Errors: Fig. 7 shows the relative depth errors that were experimentally acquired

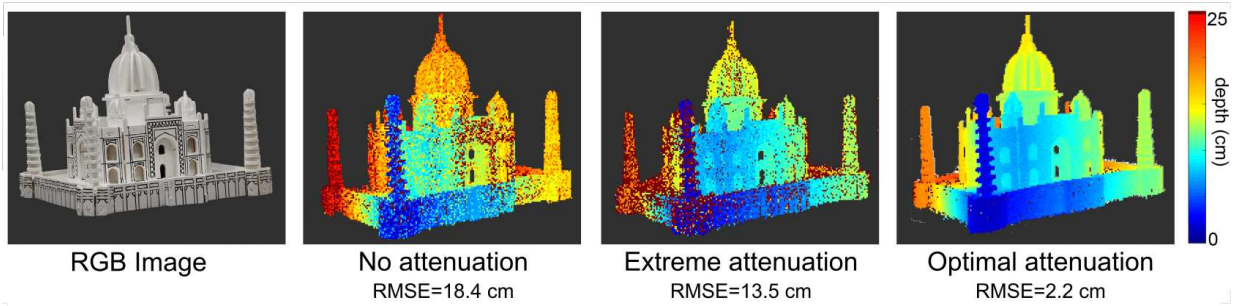

Figure 9. **Depth estimation with varying attenuation.** The average ambient illuminance of the scene was 15 000 lx. With no attenuation, most parts are affected by strong pile-up, resulting in several outliers. For extreme attenuation, large parts of the scene have very low SNR. In contrast, optimal attenuation achieves high depth estimation performance for nearly the entire object. (15 m depth offset removed.)

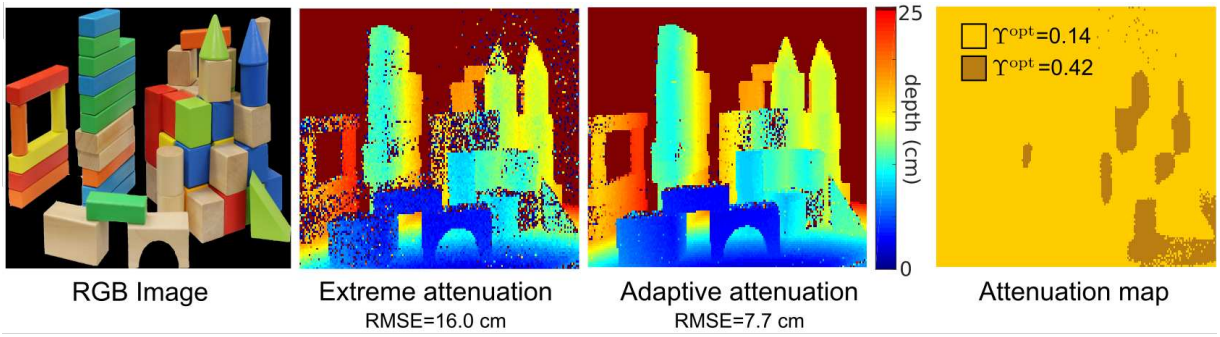

Figure 10. **Ambient-adaptive Υ^{opt}.** This scene has large ambient brightness variations, with both brightly lit regions (right) and shadows (left). Pixel-wise ambient flux estimates were used to adapt the optimal attenuation, as shown in the attenuation map. The resulting reconstruction achieves accurate estimates, both in shadows and brightly lit regions. (15 m depth offset removed.)

over a wide range of ambient and source flux levels and different attenuation factors. These experimental curves follow the same trends observed in the simulated plots of Fig. 5 and provide experimental validation for the optimal flux criterion in the presence of non-idealities like jitter and afterpulsing effects, and for a non-delta waveform.

3D Reconstructions with Point Scanning: Figs. 8 and 9 show 3D reconstruction results of objects under varying attenuation levels, acquired by raster-scanning the laser spot with a two-axis galvo-mirror system (Thorlabs GVS-012). It can be seen from the histograms in Fig. 8 (b) that extreme attenuation almost completely removes pile-up, but also reduces the signal to very low levels. In contrast, optimal attenuation has some residual pile-up, and yet, achieves approximately an order of magnitude higher depth precision as compared to extreme and no attenuation. Due to relatively uniform albedos and illumination, a single attenuation factor for the whole scene was sufficient.

Fig. 10 shows depth maps for a complex scene containing a wider range of illumination levels, albedo variations and multiple objects over a wider depth range. The optimal scheme for the "Blocks" scene adaptively chooses different attenuation factors for the parts of the scene in direct and indirect ambient light.[7] Adaptive attenuation enables depth reconstruction over a wide range of ambient flux levels.

[7] In this proof-of-concept, we acquired multiple scans at different attenuations, and stitched together the final depth map in post-processing.

8. Limitations and Future Outlook

Achieving uniform depth precision across depths: The optimal attenuation derived in this paper results in a high and relatively less skewed BRC (as shown in Fig. 4), resulting in high depth precision across the entire depth range. However, since the optimal curve has some degree of pile-up and is monotonically decreasing, later bins corresponding to larger depths still incur larger errors. It may be possible to design a time-varying attenuation scheme that gives uniform depth estimation performance.

Handling non-impulse waveforms: Our analysis assumes ideal delta waveform, as well as low source power, which allows ignoring the effect of pile-up due to the source itself. For applications where source power is comparable to ambient flux, a next step is to optimize over non-delta waveforms [13] and derive the optimal flux accordingly.

Multi-photon SPAD LiDAR: With recent improvements in detector technology, SPADs with lower dead times (tens of ns) can be realized, which enable capturing more than one photon per laser cycle. This includes multi-stop TCSPC electronics and SPADs that can be operated in the free-running mode, for which imaging models and estimators have been proposed recently [30, 14]. An interesting future direction is to derive optimal flux criterion for such multi-photon SPAD-based LiDARs.

References

[1] Supreeth Achar, Joseph R. Bartels, William L. 'Red' Whittaker, Kiriakos N. Kutulakos, and Srinivasa G. Narasimhan. Epipolar time-of-flight imaging. *ACM Trans. Graph.*, 36(4):37:1–37:8, July 2017. 2

[2] Wolfgang Becker. *Advanced time-correlated single photon counting applications*, volume 111. Springer, 2015. 2, 4, 6

[3] Maik Beer, Olaf M. Schrey, Jan F. Haase, Jennifer Ruskowski, Werner Brockherde, Bedrich J. Hosticka, and Rainer Kokozinski. Spad-based flash lidar sensor with high ambient light rejection for automotive applications. In *Quantum Sensing and Nano Electronics and Photonics XV*, volume 10540, pages 10540–10548, 2018. 2

[4] Mauro Buttafava, Gianluca Boso, Alessandro Ruggeri, Alberto Dalla Mora, and Alberto Tosi. Time-gated single-photon detection module with 110 ps transition time and up to 80 MHz repetition rate. *Review of Scientific Instruments*, 85(8):083114, 2014. 7

[5] Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Opt. Express*, 23(16):20997–21011, Aug 2015. 2

[6] Edoardo Charbon, Matt Fishburn, Richard Walker, Robert K. Henderson, and Cristiano Niclass. *SPAD-Based Sensors*, pages 11–38. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. 1

[7] P B Coates. The correction for photon 'pile-up' in the measurement of radiative lifetimes. *Journal of Physics E: Scientific Instruments*, 1(8):878, 1968. 2, 3, 4, 7

[8] Henri Dautet, Pierre Deschamps, Bruno Dion, Andrew D. MacGregor, Darleene MacSween, Robert J. McIntyre, Claude Trottier, and Paul P. Webb. Photon counting techniques with silicon avalanche photodiodes. *Appl. Opt.*, 32(21):3894–3900, Jul 1993. 2

[9] J Degnan. Impact of receiver deadtime on photon-counting slr and altimetry during daylight operations. In *16th International Workshop On Laser Ranging, Poznan Poland*, 2008. 2

[10] Daniel G Fouche. Detection and false-alarm probabilities for laser radars that use geiger-mode detectors. *Applied Optics*, 42(27):5388–5398, 2003. 2

[11] M. Gupta, Q. Yin, and S. K. Nayar. Structured light in sunlight. In *2013 IEEE International Conference on Computer Vision*, pages 545–552, Dec 2013. 2

[12] Chris M Harris and Ben K Selinger. Single-photon decay spectroscopy. ii. the pile-up problem. *Australian Journal of Chemistry*, 32(10):2111–2129, 1979. 1

[13] Felix Heide, Steven Diamond, David B. Lindell, and Gordon Wetzstein. Sub-picosecond photon-efficient 3d imaging using single-photon sensors. *Scientific Reports*, 8(1), Dec 2018. 1, 2, 4, 8

[14] Sebastian Isbaner, Narain Karedla, Daja Ruhlandt, Simon Christoph Stein, Anna Chizhik, Ingo Gregor, and Jörg Enderlein. Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging. *Optics express*, 24(9):9429–9445, 2016. 8

[15] Peter Kapusta, Michael Wahl, and Rainer Erdmann. Advanced Photon Counting Applications, Methods, Instrumentation. *Springer Series on Fluorescence*, 15, 2015. 2, 3, 4, 6

[16] Ahmed Kirmani, Dheera Venkatraman, Dongeek Shin, Andrea Colaço, Franco N. C. Wong, Jeffrey H. Shapiro, and Vivek K Goyal. First-photon imaging. *Science*, 343(6166):58–61, 2014. 2

[17] D.B. Lindell, M. O'Toole, and G. Wetzstein. Single-Photon 3D Imaging with Deep Sensor Fusion. *ACM Trans. Graph. (SIGGRAPH)*, 37(4), 2018. 2, 7

[18] Christoph Mertz, Sanjeev J Koppal, Solomon Sia, and Srinivasa Narasimhan. A low-power structured light sensor for outdoor scene reconstruction and dominant material identification. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pages 15–22. IEEE, 2012. 2

[19] Pushmeet Kohli Nathan Silberman, Derek Hoiem and Rob Fergus. Indoor segmentation and support inference from RGBD images. In *ECCV*, 2012. 7

[20] Nature Publishing Group. Lidar drives forwards. *Nature Photonics*, 12(8):441, July 2018. 1

[21] D.V. O'Connor and D. Phillips. *Time-correlated single photon counting*. Academic Press, 1984. 2

[22] Matthew O'Toole, Supreeth Achar, Srinivasa G. Narasimhan, and Kiriakos N. Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM Trans. Graph.*, 34(4):35:1–35:13, July 2015. 2

[23] M. O'Toole, F. Heide, D. B. Lindell, K. Zang, S. Diamond, and G. Wetzstein. Reconstructing transient images from single-photon sensors. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2289–2297, July 2017. 1

[24] Matthew O'Toole, David B. Lindell, and G. Wetzstein. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature*, 555:338–341, Mar 2018. 1

[25] Angus Pacala and Mark Frichtl. Optical system for collecting distance information within a field. *United States Patent 10063849*, 2018. 1

[26] Matthias Patting, Paja Reisch, Marcus Sackrow, Rhys Dowler, Marcelle Koenig, and Michael Wahl. Fluorescence decay data analysis correcting for detector pulse pile-up at very high count rates. *Optical engineering*, 57(3):031305, 2018. 1

[27] Adithya K Pediredla, Aswin C Sankaranarayanan, Mauro Buttafava, Alberto Tosi, and Ashok Veeraraghavan. Signal processing based pile-up compensation for gated single-photon avalanche diodes. *arXiv preprint arXiv:1806.07437*, 2018. 3, 4

[28] Matteo Perenzoni, Daniele Perenzoni, and David Stoppa. A 64x64-pixels digital silicon photomultiplier direct tof sensor with 100-MPhotons/s/pixel background rejection and imaging/altimeter mode with 0.14% precision up to 6 km for spacecraft navigation and landing. *IEEE Journal of Solid-State Circuits*, 52:151–160, 2017. 2

[29] J. Rapp and V. K. Goyal. A few photons among many: Unmixing signal and noise for photon-efficient active imaging. *IEEE Transactions on Computational Imaging*, 3(3):445–459, Sept 2017. 2

[30] Joshua Rapp, Yanting Ma, Robin Dawson, and Vivek K Goyal. Dead time compensation for high-flux ranging. *arXiv preprint arXiv:1810.11145*, 2018. 8

[31] D. Renker. Geiger-mode avalanche photodiodes, history, properties and problems. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 567(1):48 – 56, 2006. Proceedings of the 4th International Conference on New Developments in Photodetection. 2

[32] Daniel Scharstein and Chris Pal. Learning conditional random fields for stereo. In *IEEE Conference on Computer Vision and Pattern Recognition, 2007*, pages 1–8, 2007. 7

[33] D. E. Schwartz, E. Charbon, and K. L. Shepard. A single-photon avalanche diode array for fluorescence lifetime imaging microscopy. *IEEE Journal of Solid-State Circuits*, 43(11):2546–2557, Nov 2008. 1

[34] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro. Photon-efficient computational 3-d and reflectivity imaging with single-photon detectors. *IEEE Transactions on Computational Imaging*, 1(2):112–125, June 2015. 2

[35] John G Walker. Iterative correction for photon 'pile-up' in single-photon lifetime measurement. *Optics Communications*, 201(4-6):271–277, 2002. 2

[36] Liang Wang, Shaokun Han, Wenze Xia, and Jieyu Lei. Adaptive aperture for geiger mode avalanche photodiode flash ladar systems. *Review of Scientific Instruments*, 89(2):023105, 2018. 2