Passive Inter-Photon Imaging

Atul Ingle [a]*, Trevor Seets [a], Mauro Buttafava [b], Shantanu Gupta [a],
Alberto Tosi [b], Mohit Gupta [a]†, Andreas Velten [a]†

[a] University of Wisconsin-Madison [b] Politecnico di Milano

Abstract

Digital camera pixels measure image intensities by converting incident light energy into an analog electrical current, and then digitizing it into a fixed-width binary representation. This direct measurement method, while conceptually simple, suffers from limited dynamic range and poor performance under extreme illumination — electronic noise dominates under low illumination, and pixel full-well capacity results in saturation under bright illumination. We propose a novel intensity cue based on measuring inter-photon timing*, defined as the time delay between detection of successive photons. Based on the statistics of inter-photon times measured by a time-resolved single-photon sensor, we develop theory and algorithms for a scene brightness estimator which works over extreme dynamic range; we experimentally demonstrate imaging scenes with a dynamic range of over ten million to one. The proposed techniques, aided by the emergence of single-photon sensors such as single-photon avalanche diodes (SPADs) with picosecond timing resolution, will have implications for a wide range of imaging applications: robotics, consumer photography, astronomy, microscopy and biomedical imaging.*

1. Measuring Light from Darkness

Digital camera technology has witnessed a remarkable revolution in terms of size, cost and image quality over the past few years. Throughout this progress, however, one fundamental characteristic of a camera sensor has not changed: the way a camera pixel measures brightness. Conventional image sensor pixels manufactured with complementary metal oxide semiconductor (CMOS) and charge-coupled device (CCD) technology can be thought of as light buckets (Fig. 1(a)), which measure scene brightness in two steps: first, they collect hundreds or thousands of photons and convert the energy into an analog electrical signal (e.g. current or voltage), and then they digitize this analog quantity using high-resolution analog-to-digital converters. Conceptually, there are two main drawbacks of this image formation strategy. First, at extremely low brightness levels, noise in the pixel electronics dominates resulting in poor signal-to-noise-ratio (SNR). Second, since each pixel bucket has a fixed maximum capacity, bright regions in the scene cause the pixels to saturate and subsequent incident photons do not get recorded.

In this paper, we explore a different approach for measuring image intensities. Instead of estimating intensities directly from the number of photons incident on a pixel, we propose a novel intensity cue based on *inter-photon timing*, defined as the time delay between detection of successive photons. Intuitively, as the brightness increases, the *time-of-darkness* between consecutive photon detections decreases. By modeling the statistics of photon arrivals, we derive a theoretical expressions that relates the average inter-photon delay and the incident flux. The key observation is that because photon arrivals are stochastic, the average inter-photon time decreases asymptotically as the incident flux increases. Using this novel temporal intensity cue, we design algorithms to estimate brightness from as few as one photon timestamp per pixel to extremely high brightness, beyond the saturation limit of conventional sensors.

How to Measure Inter-Photon Timing? The inter-photon timing intensity cue and the resulting brightness estimators can achieve extremely high dynamic range. A natural question to ask then is: How does one measure the inter-photon timing? Conventional CMOS sensor pixels do not have the ability to measure time delays between individual photons at the timing granularity needed for estimating intensities with high precision. Fortunately, there is an emerging class of sensors called single-photon avalanche diodes (SPADs) [11, 7], that can not only detect individual photons, but also precisely time-tag each captured photon with picosecond resolution.

Emergence of Single-Photon Sensors: SPADs are naturally suited for imaging in low illumination conditions, and thus, are fast becoming the sensors of choice for applications that require extreme sensitivity to photons together with fine-grained temporal information: single-photon 3D time-of-flight imaging [52, 34, 46, 45, 33], transient imaging [50, 49], non-line-of-sight imaging [35, 19], and flu-

†Equal contribution.

This research was supported in part by DARPA HR0011-16-C-0025, DoE NNSA DE-NA0003921 [1], NSF GRFP DGE-1747503, NSF CAREER 1846884 and 1943149, and Wisconsin Alumni Research Foundation.

* Email: ingle@uwalumni.com

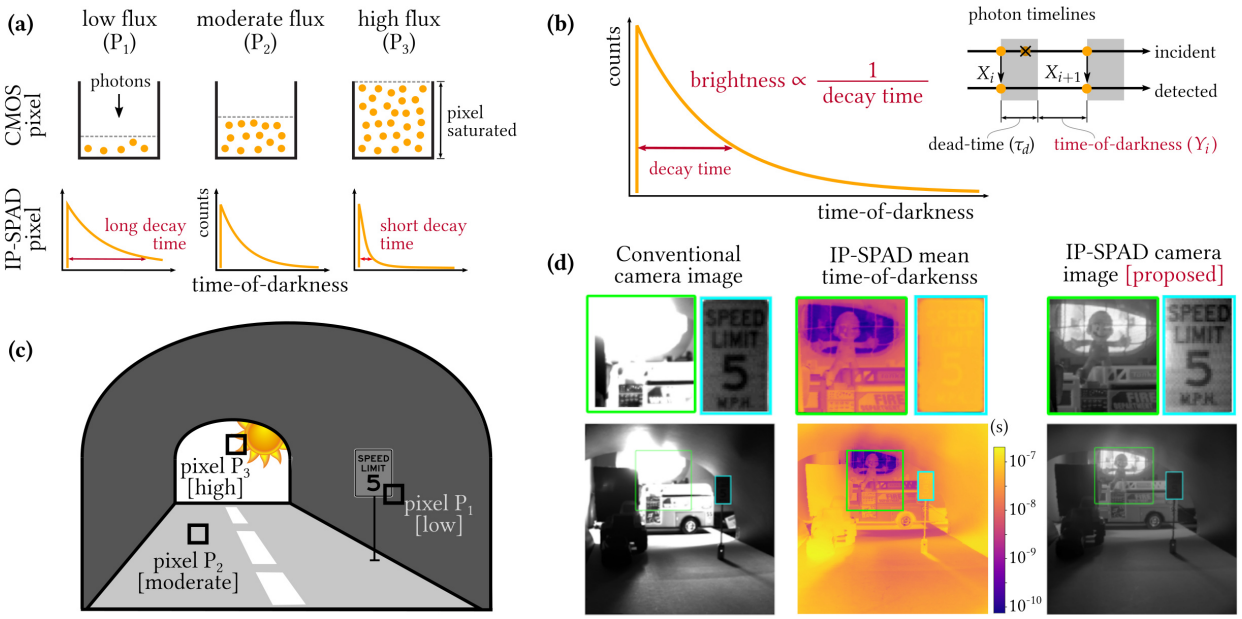


Figure 1: **Passive imaging with an inter-photon single-photon avalanche diode (IP-SPAD):** (a) A conventional image sensor pixel estimates scene brightness using a well-filling mechanism; the well has a finite capacity and saturates for very high brightness levels. (b) An IP-SPAD measures scene brightness from inter-photon timing measurements that follow Poisson statistics. The higher the brightness, the smaller the inter-photon time, the faster the decay rate of the inter-photon histogram. By capturing photon timing information with very high precision, this estimator can provide scene brightness estimates well beyond the saturation limit of conventional pixels. (c) A representative extreme dynamic range scene of a tunnel with three different flux levels (low, moderate and high) shown for illustration. (d) Experimental results from our hardware prototype comparing a conventional CMOS camera image and an image obtained from our IP-SPAD prototype.

orescence microscopy [43]. While these applications use SPADs in active imaging setups in synchronization with an illumination source such as a pulsed laser, recently these sensors have been explored as passive, general-purpose imaging devices for high-speed and high-dynamic range photography [4, 25, 36]. In particular, it was shown that SPADs can be used to measure incident flux while operating as passive, free-running pixels (PF-SPAD imaging) [25]. The dynamic range of the resulting measurements, although higher than conventional pixels (that rely on a finite-depth well filling light detection method like CCD and CMOS sensors), is inherently limited due to *quantization* stemming from the discrete nature of photon counts.

Intensity from Inter-Photon Timing: Our key idea is that it is possible to circumvent the limitations of counts-based photon flux estimation by exploiting photon timing information from a SPAD. The additional time-dimension is a rich source of information that is inaccessible to conventional photon count-based methods. We derive a scene brightness estimator that relies on the decay time statistics of the inter-photon times captured by a SPAD sensor as shown in Fig. 1(b). We call our image sensing method *inter-photon SPAD (IP-SPAD)* imaging. An IP-SPAD pixel captures the decay time distribution which gets narrower with increasing brightness. As shown in Fig. 1(d), the measurements can be summarized in terms of the mean time-of-darkness, which can then be used to estimate incident flux.

Unlike a photon-counting PF-SPAD pixel whose measurements are inherently discrete, an IP-SPAD measures decay times as floating point values, capturing information at much finer granularity than integer-valued counts, thus enabling measurement of extremely high flux values. In practice, the dynamic range of an IP-SPAD is limited only by the precision of the floating point representation used for measuring the time-of-darkness between consecutive photons. Coupled with the sensitivity of SPADs to individual photons and lower noise compared to conventional sensors, the proposed approach, for the first time, achieves ultra high-dynamic range. We experimentally demonstrate a dynamic range of over ten million to one, simultaneously imaging extremely dark (pixels P_1 and P_2 inside the tunnel in Fig. 1(c)) as well as very bright scene regions (pixel P_3 outside the tunnel in Fig. 1(c)).

2. Related Work

High-Dynamic-Range Imaging: Conventional high-dynamic-range (HDR) imaging techniques using CMOS

image sensors use variable exposures to capture scenes with extreme dynamic range. The most common method called exposure bracketing [21, 22] captures multiple images with different exposure times; shorter exposures reliably capture bright pixels in the scene avoiding saturation, while longer exposures capture darker pixels while avoiding photon noise. Another technique involves use of neutral density (ND) filters of varying densities resulting in a tradeoff between spatial resolution and dynamic range [40]. ND filters reduce overall sensitivity to avoid saturation, at the cost of making darker scene pixels noisier. In contrast, an IP-SPAD captures scene intensities in a different way by relying on the non-linear reciprocal relationship between inter-photon timing and scene brightness. This gives extreme dynamic range in a single capture.

Passive Imaging with Photon-Counting Sensors: Previous work on passive imaging with photon counting sensors relies on two sensor technologies—SPADs and quanta-image sensors (QISs) [18]. A QIS has single-photon sensitivity but much lower time resolution than a SPAD pixel. On the other hand, QIS pixels can be designed with much smaller pixel pitch compared to SPAD pixels, allowing spatial averaging to further improve dynamic range while still maintaining high spatial resolution [36]. SPAD-based high-dynamic range schemes provide lower spatial resolution than the QIS-based counterparts [16], although, recently, megapixel SPAD arrays capable of passive photon counting have also be developed [39]. Previous work [25] has shown that passive free-running SPADs can potentially provide several orders of magnitude improved dynamic range compared to conventional CMOS image sensor pixel. The present work exploits the precise *timing information*, in addition to photon counts, measured by a free-running SPAD sensor. An IP-SPAD can image scenes with even higher dynamic range than the counts-based PF-SPAD method.

Methods Relying on Photon Timing: The idea of using timing information to increase dynamic range has been explored before for conventional CMOS image sensor pixels. A saturated CMOS pixel's output is simply a constant and meaningless, but if the time taken to reach saturation is also available [13], it provides information about scene brightness, because a brighter scene point will reach saturation more quickly (on average) than a dimmer scene point. The idea of using photon timing information for HDR has also been discussed before but the dynamic range improvements were limited by the low timing resolution of the pixels [53, 31] at which point, the photon timing provides no additional information over photon counts.

Methods Relying on Non-linear Sensor Response: Logarithmic image sensors include additional pixel electronics that apply log-compression to capture a large dynamic range. These pixels are difficult to calibrate and require additional pixel electronics compared to conventional CMOS image sensor pixels [27]. A modulo-camera [54] allows a conventional CMOS pixel output to wrap around after saturation. It requires additional in-pixel computation involving an iterative algorithm that unwraps the modulo-compression to reconstruct the high-dynamic-range scene. In contrast, our timing-based HDR flux estimator is a closed-form expression that can be computed using simple arithmetic operations. Although our method also requires additional in-pixel electronics to capture high-resolution timing information, recent trends in SPAD technology indicate that such arrays can be manufactured cheaply and at scale using CMOS fabrication techniques [24, 23].

Active Imaging Methods: Photon timing information captured by a SPAD sensor has been exploited for various active imaging applications like transient imaging [41], fluorescence lifetime microscopy [7], 3D imaging LiDAR [30, 20] and non-line-of-sight imaging [29, 9]. Active methods capture photon timing information with respect to a synchronized light source like a pulsed laser that illuminates the scene. In contrast, we operate the SPAD asynchronously and measure the time between successive photons in a passive imaging setting where the scene is only illuminated by ambient light.

3. Image Formation with Inter-Photon Timing

3.1. Flux Estimator

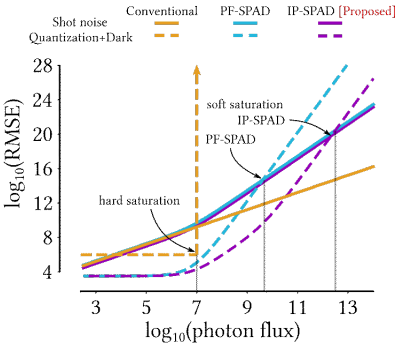
Consider a single IP-SPAD pixel passively capturing photons over a fixed exposure time T from a scene point[1] with true photon flux of Φ photons per second. After each photon detection event, the IP-SPAD pixel goes blind for a fixed duration τ_{d} called the dead-time. During this dead-time, the pixel is reset and the pixel's time-to-digital converter (TDC) circuit stores a picosecond resolution timestamp of the most recent photon detection time, and also increments the total photon count. This process is repeated until the end of the exposure time T. Let $N_T \geq 2$ denote the total number of photons detected by the IP-SPAD pixel during its fixed exposure time, and let X_i $(1 \leq i \leq N_T)$ denote the timestamp of the i^{th} photon detection. The measured inter-photon times between successive photons is defined as $Y_i := X_{i+1} - X_i - \tau_{\mathrm{d}}$ (for $1 \leq i \leq N_T - 1$). Note that Y_i's follow an exponential distribution. It is tempting to derive a closed-from maximum likelihood photon flux estimator $\widehat{\Phi}$ for the true flux Φ using the log-likelihood function of the

[1] We assume that there is no scene or camera motion so that the flux Φ stays constant over the exposure time T.

Sources of Noise in Flux Estimation: Theoretical Expressions

Noise Sources	Conventional		PF-SPAD		IP-SPAD [Proposed]	
	Bias	Variance	Bias	Variance	Bias	Variance
Shot noise	-	$\frac{\Phi}{qT}$	-	$\frac{\Phi(1+q\Phi\tau_d)}{qT}$	-	$\frac{\Phi(1+q\Phi\tau_d)}{qT}$
Dark noise	-	$\frac{\sigma_r^2}{q^2T^2}$	Φ_{dark}	$\frac{\Phi_{dark}(1+q\Phi_{dark}\tau_d)}{qT}$	Φ_{dark}	$\frac{\Phi_{dark}(1+q\Phi_{dark}\tau_d)}{qT}$
Quantization noise	-	≈0, if not saturated ∞, if saturated	-	$\frac{(1+q\Phi\tau_d)^4}{12q^2T^2}$	-	$\frac{(1+q\Phi(\tau_d+\Delta))^2(1+q\Phi\Delta)^2}{12q^2T^2}$

(a)



(b)

Figure 2: **Comparison of noise sources in different image sensor pixels:** (a) Theoretical expressions for the three main sources of noise affecting a conventional pixel, PF-SPAD pixel [25] and the proposed IP-SPAD pixel are summarized in this table. Note that the IP-SPAD's sources of noise are similar to a PF-SPAD except for quantization noise. (b) The expressions in (a) are plotted for the case of $T = 5\,\text{ms}$, $q = 100\%$, $\sigma_r = 5e^-$, $\Phi_{\text{dark}} = 10$ photons/second, $\tau_d = 150\,\text{ns}$, $\Delta = 200\,\text{ps}$. The conventional sensor's saturation capacity is set at 34,000 e^- which matches the maximum possible SPAD counts of $\lceil T/\tau_d \rceil$. Observe that the IP-SPAD soft-saturation point is at a much higher flux level than the PF-SPAD.

measured inter-photon times Y_i:

$$\log l(q\Phi; Y_1, \ldots, Y_{N_T-1}) = \log\left(\prod_{i=1}^{N_T-1} q\Phi\, e^{-q\Phi Y_i}\right)$$
$$= -q\Phi\left(\sum_{n=1}^{N_T-1} Y_i\right) + (N_T - 1)\log q\Phi, \quad (1)$$

where $0< q <1$ is the quantum efficiency of the IP-SPAD pixel. The maximum likelihood estimate $\widehat{\Phi}$ of the true photon flux is computed by setting the derivative of Eq. (1) to zero and solving for Φ:

$$\boxed{\widehat{\Phi} = \frac{1}{q}\frac{N_T - 1}{X_{N_T} - X_1 - (N_T - 1)\tau_d}.} \quad (2)$$

Although the above proof sketch captures the intuition of our flux estimator, it leaves out two details. First, the total number of photons N_T is itself a random variable. Second, the times of capture of future photons are constrained by the timestamps of preceding photon arrivals because we operate in a finite exposure time T. The sequence of timestamps Y_i cannot be treated as independent and identically distributed. The conditional distribution of the i^{th} inter-photon time conditioned on the previous inter-photon times is given by:

$$p_{Y_i|Y_1,\ldots,Y_{i-1}}(t) = \begin{cases} q\Phi e^{-q\Phi t} & 0 < Y_i < T_i \\ e^{-q\Phi T_i}\delta(t - T_i) & Y_i = T_i \\ 0 & \text{otherwise.} \end{cases}$$

Here $\delta(\cdot)$ is the Dirac delta function. The T_i's model the shrinking effective exposure times for subsequent photon detections. $T_1 = T$ and for $i > 1$, $T_i = \max(0, T_{i-1} - Y_{i-1} - \tau_d)$. The log-likelihood function can now be written as:

$$\log l(q\Phi; Y_1, \ldots, Y_L) = \log\left(\prod_{i=1}^{\lceil T/\tau_d \rceil} p_{Y_i|Y_1,\ldots,Y_{i-1}}(t)\right)$$
$$= -q\Phi \max\left(\sum_{i=1}^{N_T} Y_i, T - N_T\tau_d\right) + N_T \log q\Phi.$$

As shown in Supplementary Note 1 this likelihood function also leads to the flux estimator given in Eq. (2)

We make the following key observations about the IP-SPAD flux estimator. First, note that the estimator is only a function of the first and the last photon timestamps, the exact times of capture of the intermediate photons do not provide additional information.[2] This is because photon arrivals follow Poisson statistics: the time until the next photon arrival from the end of the previous dead-time is independent of all preceding photon arrivals. Secondly, we note that the denominator in Eq. (2) is simply the total time the IP-SPAD spends waiting for the next photon to be captured while not in dead-time. Intuitively, if the photon flux were extremely high, we will expect to see a photon immediately after every dead-time duration ends, implying the denominator in Eq. (2) approaches zero, hence $\widehat{\Phi} \to \infty$.

3.2. Sources of Noise

Although, theoretically, the IP-SPAD scene brightness estimator in Eq. (2) can recover the entire range of incident photon flux levels, including very low and very high

[2] As we show later in our hardware implementation, in practice, it is useful to capture intermediate photon timestamps as they allow us to calibrate for various pixel non-idealities.

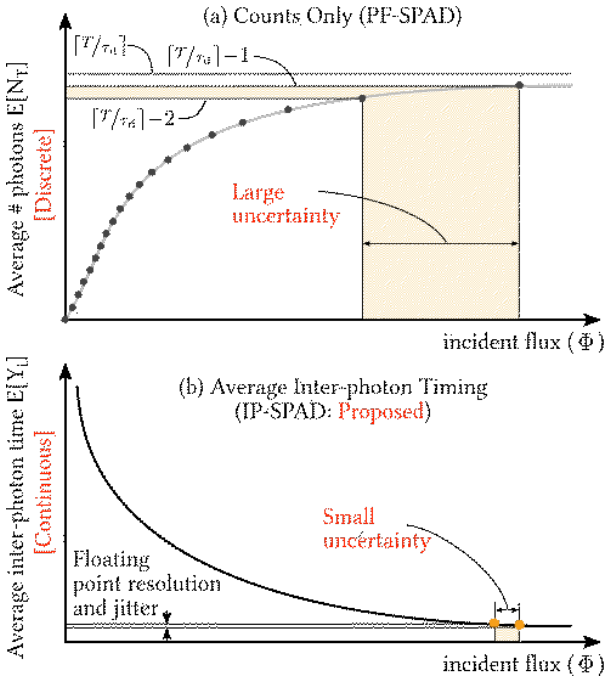


Figure 3: **Advantage of using photon timing over photon counts:** (a) Photon counts are inherently discrete. At high flux levels, even a small ± 1 change in photon counts corresponds to a large flux uncertainty. (b) Inter-photon timing is inherently continuous. This leads to smaller uncertainty at high flux levels. The uncertainty depends on jitter and floating point resolution of the timing electronics.

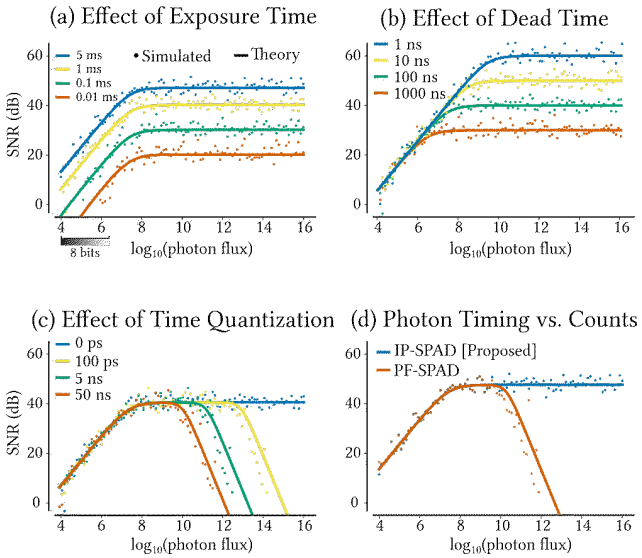


Figure 4: **Effect of various IP-SPAD parameters on SNR:** We vary different parameters to see the effect on SNR. The solid lines are theoretical SNR curves while each dot represents a SNR average from 10 Monte Carlo simulations. Unless otherwise noted the parameters used are $T = 1$ ms, $\tau_d = 100$ ns, $q = 0.4$, and $\Delta = 0$. (a) As exposure time increases, SNR increases at all brightness levels. (b) Decreasing the dead-time increases the maximum achievable SNR, but provides little benefit in low flux. (c) Coarser time quantization causes SNR drop-off at high flux values. (d) Our IP-SPAD flux estimator outperforms a counts-only (PF-SPAD) flux estimator [25] at high flux levels.

flux values, in practice, the accuracy is limited by various sources of noise. To assess the performance of this estimator, we use a signal-to-noise ratio (SNR) metric defined as [51, 25]:

$$\mathrm{SNR}(\Phi) = 10 \log_{10} \left(\frac{\Phi^2}{\mathbf{E}[(\Phi - \widehat{\Phi})^2]} \right) \tag{3}$$

Note that the denominator in Eq. (3) is the mean-squared-error of the estimator $\widehat{\Phi}$, and is equal to the sum of the bias-squared terms and variances of the different sources of noise. The *dynamic range* (DR) of the sensor is defined as the range of brightness levels for which the SNR stays above a minimum specified threshold. At extremely low flux levels, the dynamic range is limited due to IP-SPAD dark counts which causes spurious photon counts even when no light is incident on the pixel. This introduces a bias in $\widehat{\Phi}$. Since photon arrivals are fundamentally stochastic (due to the quantum nature of light), the estimator also suffers from Poisson noise which introduces a non-zero variance term. Finally, at high enough flux levels, the time discretization Δ used for recording timestamps with the IP-SPAD pixel limits the maximum usable photon flux at which the pixel can operate. Fig. 2(a) shows the theoretical expression for bias and variance introduced by shot noise, quantization noise and dark noise in an IP-SPAD pixel along with corresponding expressions for a conventional image sensor pixel and a PF-SPAD pixel. Fig. 2(b) shows example plots for these theoretical expressions. For realistic values of Δ in the 100's of picoseconds range, the IP-SPAD pixel has a smaller quantization noise term that allows reliable brightness estimation at much higher flux levels than a PF-SPAD pixel. (See Supplementary Note 2).

Quantization Noise in PF-SPAD vs. IP-SPAD: Conventional pixels are affected by quantization in low flux and hard saturation (full-well capacity) limit in high flux. In contrast, a PF-SPAD pixel that only uses photon counts is affected by quantization noise *at extremely high flux levels due to soft-saturation* [25]. This behavior is unique to SPADs and is quite different from conventional sensors. A counts-only PF-SPAD pixel can measure at most $\lceil T/\tau_d \rceil$ photons where T is the exposure time and τ_d is the dead-time [25]. Due to a non-linear response curve, as shown in Fig. 3(a), a small change of ± 1 count maps to a large range of flux values. Due to the inherently discrete nature of photon counts, even a small fluctuation (due to shot noise

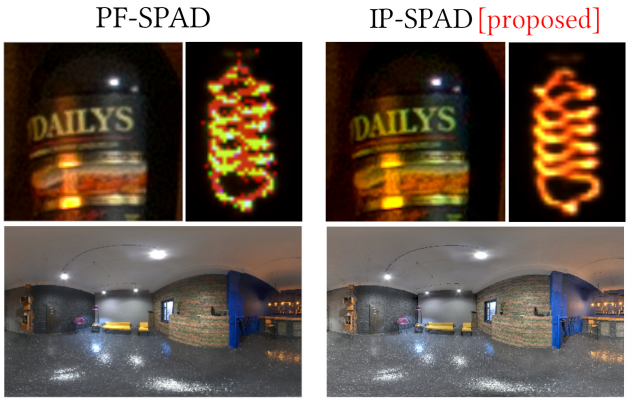


Figure 5: **Simulated HDR scene captured with a PF-SPAD (counts only) vs. IP-SPAD (inter-photon timing):** (a) Although a PF-SPAD can capture this extreme dynamic range scene in a single $5\,\mathrm{ms}$ exposure, extremely bright pixels such as the bulb filament that are beyond the soft-saturation limit appear noisy. (b) An IP-SPAD camera captures both dark and bright regions in a single exposure, including fine details around the bright bulb filament. In both cases, we set the SPAD pixel's quantum efficiency to 0.4, dead-time to $150\,\mathrm{ns}$ and an exposure time of $5\,\mathrm{ms}$. The IP-SPAD has a time resolution of $\Delta = 200\,\mathrm{ps}$. (Original image from HDRIHaven.com)

or jitter) can cause a large uncertainty in the estimated flux.

The proposed IP-SPAD flux estimator uses timing information which is inherently continuous. Even at extremely high flux levels, photon arrivals are random and due to small random fluctuations, the time interval between the first and last photon ($X_{N_T} - X_1$) is not exactly equal to T. Fig. 3(b) shows the intuition for why fine-grained inter-photon measurements at high flux levels can enable flux estimation with a smaller uncertainty than counts alone. In practice, the improvement in dynamic range compared to a PF-SPAD depends on the time resolution, which is limited by hardware constraints like floating point precision of the TDC electronics and timing jitter of the SPAD pixel. Simulations in Fig. 4 suggest that even with a $100\,\mathrm{ps}$ time resolution the 20-dB dynamic range improves by 2 orders of magnitude over using counts alone.

Single-Pixel Simulations: We verify our theoretical SNR expression using single-pixel Monte Carlo simulations of a single IP-SPAD pixel. For a fixed set of parameters we run 10 simulations of an IP-SPAD at 100 different flux levels ranging $10^4 - 10^{16}$ photons per second. Fig. 4 shows the effect of various pixel parameters on the SNR. The overall SNR can be increased by either increasing the exposure time T or decreasing the dead-time τ_d; both enable the IP-SPAD pixel to capture more total photons. The maximum achievable SNR is theoretically equal to $10\log_{10}(T/\tau_\mathrm{d})$. The IP-SPAD SNR degrades at high flux levels due because photon timestamps cannot be captured with infinite resolution. A larger floating point quantization bin size Δ increases the uncertainty in photon timestamps. If the time bin is large enough, there is no advantage in using the timestamp-based brightness estimator and the performance reverts to a counts-based flux estimator [4, 25].

4. Results

4.1. Simulation Results

Simulated RGB Image Comparisons: Fig. 5 shows a simulated HDR scene with extreme brightness variations between the dark text and bright bulb filament. We use a $5\,\mathrm{ms}$ exposure time for this simulation. The PF-SPAD and IP-SPAD both use pixels with $q = 0.4$ and $\tau_\mathrm{d} = 150\,\mathrm{ns}$. The PF-SPAD only captures photon counts whereas the IP-SPAD captures counts and timestamps with a resolution of $\Delta = 200\,\mathrm{ps}$. Notice that the extremely bright pixels on the bulb filament appear noisy in the PF-SPAD image. This degradation in SNR at high flux levels is due to its soft-saturation phenomenon. The IP-SPAD, on the other hand, captures the dark text and noise-free details in the bright bulb filament in a single exposure. Please see supplement for additional comparisons with a conventional camera.

4.2. Single-Pixel IP-SPAD Hardware Prototype

Our single-pixel IP-SPAD prototype is a modified version of a fast-gated SPAD module [8]. Conventional dead-time control circuits for a SPAD rely on digital timers that have a coarse time-quantization limited by the digital clock frequency and cannot be used for implementing an IP-SPAD. We circumvent this limitation by using coaxial cables and low-jitter voltage comparators to generate "analog delays" that enable precise control of the dead-time with jitter limited to within a few ps. We used a $20\,\mathrm{m}$ long co-axial cable to get a dead-time of $110\,\mathrm{ns}$. The measured dead-time jitter was $\sim 200\,\mathrm{ps}$. This is an improvement over previous PF-SPAD implementations [25] that relied on a digital timer circuit whose time resolution was limited to $\sim 6\,\mathrm{ns}$.

The IP-SPAD pixel is mounted on two translation stages that raster scan the image plane of a variable focal length lens. The exposure time per pixel position depends on the translation speed along each scan-line. We capture 400×400 images with an exposure time of $5\,\mathrm{ms}$ per pixel position. The total capture takes ~ 15 minutes. Photon timestamps are captured with a $1\,\mathrm{ps}$ time binning using a time-correlated single-photon counting (TCSPC) system (PicoQuant HydraHarp400). A monochrome camera (PointGrey Technologies GS3-U3-23S6M-C) placed next to the SPAD captures conventional camera images for comparison. The setup is arranged carefully to obtain approxi-

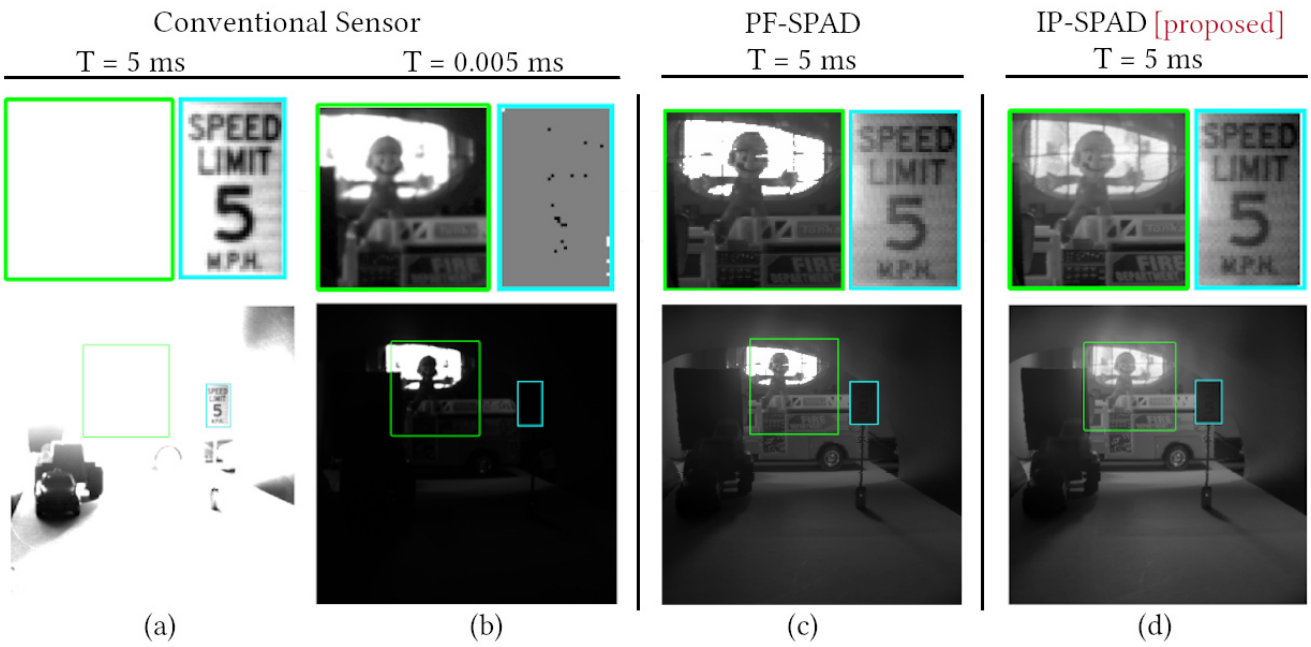


Figure 6: **Experimental "Tunnel" scene:** (a-b) Images from a conventional sensor with long and short exposure times. Notice that both the speed limit sign and the toy figure cannot be captured simultaneously with a single exposure. Objects outside the tunnel appear saturated even with the shortest exposure time possible with our CMOS camera. (c) A PF-SPAD [25] only uses photon counts when estimating flux. Although it captures much higher dynamic range than the conventional CMOS camera, the bright pixels near the halogen lamp appear saturated. (d) Our IP-SPAD single-pixel hardware prototype captures both the dark and the extremely bright regions with a single exposure. Observe that the fine details within the halogen lamp are visible.

mately the same field of view and photon flux per pixel for both the IP-SPAD and CMOS camera pixels.

4.3. Hardware Experiment Results

HDR Imaging: Fig. 6 shows an experiment result using our single-pixel raster-scanning hardware prototype. The "Tunnel" scene contains dark objects like a speed limit sign inside the tunnel and an extremely bright region outside the tunnel from a halogen lamp. This scene has a dynamic range of over 10^7:1. The conventional CMOS camera (Fig. 6(a-b)), requires multiple exposures to cover this dynamic range. Even with the shortest possible exposure time of $0.005\,\mathrm{ms}$, the halogen lamp appears saturated. Our IP-SPAD prototype captures details of both the dark regions (text on the speed limit sign) simultaneously with the bright pixels (outline of halogen lamp tube) in a single exposure. Fig. 6(c) and (d) shows experimental comparison between a PF-SPAD (counts-only) image [25] and the proposed IP-SPAD image that uses photon timestamps. Observe that in extremely high flux levels (in pixels around the halogen lamp) the PF-SPAD flux estimator fails due to the inherent quantization limitation of photon counts. The IP-SPAD preserves details in these extremely bright regions, like the shape of the halogen lamp tube and the metal cage on the lamp.

Hardware Limitations: The IP-SPAD pixel does not exit the dead-time duration instantaneously and in practice it takes around $100\,\mathrm{ps}$ to transition into a fully-on state. Representative histograms for four different locations in the experiment tunnel scene are shown in Fig. 7. Observe that at lower flux levels (pixels [P1] and [P2]) the inter-photon histograms follow an exponential distribution as predicted by the Poisson model for photon arrival statistics. However, at pixels with extremely high brightness levels (pixels [P3] and [P4] on the halogen lamp), the histograms have a rising edge denoting the transition phase when the pixel turns on after the end of the previous dead-time. We also found that in practice the dead-time is not constant and exhibits a drift over time (especially at high flux values) due to internal heating. Such non-idealities, if not accounted for, can cause uncertainty in photon timestamp measurements, and limit the usability of our flux estimator in the high photon flux regime. Since we capture timestamps for every photon detected in a fixed exposure time, it is possible to correct these non-idealities in post-processing by estimating the true dead-time and rise-time from these inter-photon timing histograms. See Supplementary Note 5 for details.

IP-SPAD Imaging with Low Photon Counts: The results so far show that precise photon timestamps from an IP-

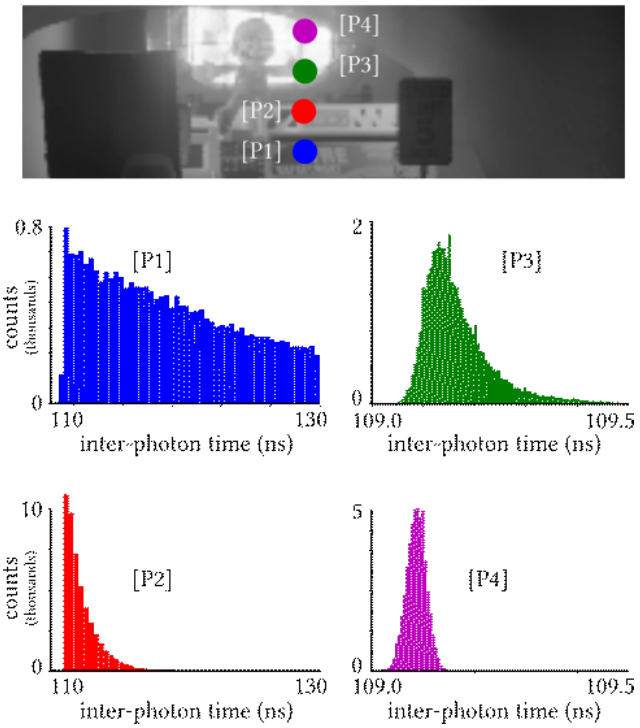


Figure 7: **Rise-time Non-ideality in Measured IP-SPAD Histograms:** We show four inter-photon histograms for pixels in the experimental "Tunnel" scene. The histograms of [P1] and [P2] have an ideal exponentially decaying shape. However, at the brighter points, [P3] and [P4], the inter-photon histograms deviate from an ideal exponential shape. This is because the IP-SPAD pixels requires $\sim$ 100 ps rise time to re-activate after the end of the previous dead-time.

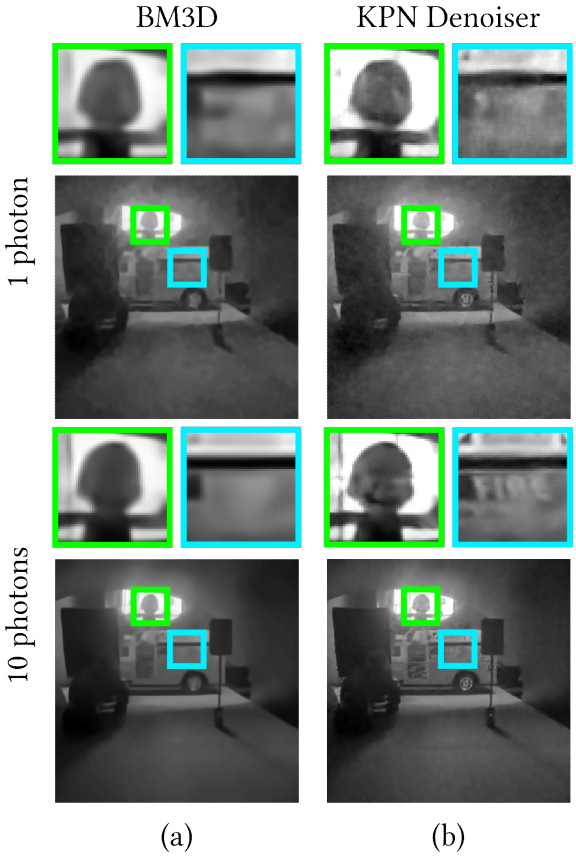


Figure 8: **IP-SPAD Imaging in Low Photon Count Regime:** This figure shows IP-SPAD images captured with very few photons and denoised with two different methods: (a) an off-the-shelf BM3D denoiser, and (b) a DNN denoiser based on a kernel prediction network architecture. Details like the text on the fire-truck are visible with as few as 10 photons per pixel.

SPAD pixel enables imaging at extremely high photon flux levels. We now show that it is also possible to leverage timing information when the IP-SPAD pixel captures very few photons per pixel. We simulate the low photon count regime by keeping the first few photons and discarding the remaining photon timestamps for each pixel in the experimental "Tunnel" scene. Fig. 8 shows IP-SPAD images captured with as few as 1 and 10 photons per pixel and denoised using an off-the-shelf BM3D denoiser and a deep neural network denoiser that uses a kernel prediction network (KPN) architecture [38]. We can recover intensity images with just one photon timestamp per pixel using real data captured by our IP-SPAD hardware prototype. Quite remarkably, with as few as 10 photons per pixel, image details such as facial features and text on the fire truck are discernible. Please see Supplementary Note 3 for details about the KPN denoiser and Supplementary Note 6 for additional experimental results and comparisons with other denoising algorithms.

5. Future Outlook

The analysis and experimental proof-of-concept shown in this paper were restricted to a single IP-SPAD pixel. Recent advances in CMOS SPAD technology that rely on 3D stacking [23] can enable larger arrays of SPAD pixels for passive imaging. This will introduce additional design challenges and noise sources not discussed here. In Supplementary Note 7 we show some pixel architectures for an IP-SPAD array that could be implemented in the future.

Arrays of single-photon image sensor pixels are being increasingly used for 2D intensity imaging and 3D depth sensing [52, 32, 37] in commercial and consumer devices. When combined with recent advances in high-time-resolution SPAD sensor hardware, the methods developed in this paper can enable extreme imaging applications across various applications including consumer photography, vision sensors for autonomous driving and robotics, and biomedical optical imaging.

References

[1] **U.S. Department of Energy (Disclaimer):** This work was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third partys use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, its contractors or subcontractors. 1

[2] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017. 7

[3] F J Anscombe. The Transformation of Poisson, Binomial and Negative-Binomial Data. *Biometrika*, 35(3-4):246 – 254, Dec. 1948. 7

[4] Ivan Michel Antolovic, Claudio Bruschini, and Edoardo Charbon. Dynamic range extension for photon counting arrays. *Optics Express*, 26(17):22234, aug 2018. 2, 6, 5

[5] Steve Bako, Thijs Vogels, Brian Mcwilliams, Mark Meyer, Jan Novák, Alex Harvill, Pradeep Sen, Tony Derose, and Fabrice Rousselle. Kernel-predicting convolutional networks for denoising monte carlo renderings. *ACM Trans. Graph.*, 36(4), July 2017. 7, 8

[6] Danilo Bronzi, Federica Villa, Simone Tisa, Alberto Tosi, and Franco Zappa. Spad figures of merit for photon-counting, photon-timing, and imaging applications: a review. *IEEE Sensors Journal*, 16(1):3–12, 2015. 16

[7] Claudio Bruschini, Harald Homulle, Ivan Michel Antolovic, Samuel Burri, and Edoardo Charbon. Single-photon avalanche diode imagers in biophotonics: review and outlook. *Light: Science & Applications*, 8(1):1–28, 2019. 1, 3

[8] 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. 6, 11

[9] Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Optics express*, 23(16):20997–21011, 2015. 3

[10] Edoardo Charbon, Claudio Bruschini, and Myung-Jae Lee. 3d-stacked CMOS SPAD image sensors: Technology and applications. In *2018 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. IEEE, dec 2018. 16, 17

[11] Sergio Cova, Massimo Ghioni, Andrea Lacaita, Carlo Samori, and Franco Zappa. Avalanche photodiodes and quenching circuits for single-photon detection. *Applied optics*, 35(12):1956–1976, 1996. 1

[12] S. Cova, A. Lacaita, and G. Ripamonti. Trapping phenomena in avalanche photodiodes on nanosecond scale. *IEEE Electron Device Letters*, 12(12):685–687, 1991. 11

[13] Eugenio Culurciello, Ralph Etienne-Cummings, and Kwabena A Boahen. A biomorphic digital image sensor. *IEEE Journal of Solid-State Circuits*, 38(2):281–294, 2003. 3

[14] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising with block-matching and 3d filtering. *Proceedings of SPIE - The International Society for Optical Engineering*, 6064:354–365, 02 2006. 8

[15] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007. 7

[16] Neale AW Dutton, Tarek Al Abbas, Istvan Gyongy, Francescopaolo Mattioli Della Rocca, and Robert K Henderson. High dynamic range imaging at the quantum limit with single photon avalanche diode-based image sensors. *Sensors*, 18(4):1166, 2018. 3

[17] M. Ghioni, A. Gulinatti, I. Rech, F. Zappa, and S. Cova. Progress in silicon single-photon avalanche diodes. *IEEE Journal of Selected Topics in Quantum Electronics*, 13(4):852–862, 2007. 17

[18] Abhiram Gnanasambandam, Omar Elgendy, Jiaju Ma, and Stanley H Chan. Megapixel photon-counting color imaging using quanta image sensor. *Optics express*, 27(12):17298–17310, 2019. 3

[19] Javier Grau Chopite, Matthias B Hullin, Michael Wand, and Julian Iseringhausen. Deep non-line-of-sight reconstruction. *arXiv*, pages arXiv–2001, 2020. 1

[20] Anant Gupta, Atul Ingle, and Mohit Gupta. Asynchronous single-photon 3d imaging. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 7909–7918, 2019. 3

[21] Mohit Gupta, Daisuke Iso, and Shree K. Nayar. Fibonacci exposure bracketing for high dynamic range imaging. In *2013 IEEE International Conference on Computer Vision*. IEEE, dec 2013. 3

[22] S. W. Hasinoff, F. Durand, and W. T. Freeman. Noise-optimal capture for high dynamic range photography. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 553–560, June 2010. 3

[23] Robert K. Henderson, Nick Johnston, Sam W. Hutchings, Istvan Gyongy, Tarek Al Abbas, Neale Dutton, Max Tyler, Susan Chan, and Jonathan Leach. 5.7 a 256×256 40nm/90nm CMOS 3d-stacked 120db dynamic-range reconfigurable time-resolved SPAD imager. In *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*. IEEE, feb 2019. 3, 8, 16, 17

[24] Robert K. Henderson, Nick Johnston, Francescopaolo Mattioli Della Rocca, Haochang Chen, David Day-Uei Li, Graham Hungerford, Richard Hirsch, David Mcloskey, Philip Yip, and David J. S. Birch. A 192x128 time correlated SPAD

image sensor in 40-nm CMOS technology. *IEEE Journal of Solid-State Circuits*, 54(7):1907–1916, jul 2019. 3, 16

[25] Atul Ingle, Andreas Velten, and Mohit Gupta. High flux passive imaging with single-photon sensors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6760–6769, 2019. 2, 3, 4, 5, 6, 7

[26] Steven D. Johnson, Paul-Antoine Moreau, Thomas Gregory, and Miles J. Padgett. How many photons does it take to form an image? *Applied Physics Letters*, 116(26):260504, jun 2020. 6

[27] Spyros Kavadias, Bart Dierickx, Danny Scheffer, Andre Alaerts, Dirk Uwaerts, and Jan Bogaerts. A logarithmic response cmos image sensor with on-chip calibration. *IEEE Journal of Solid-state circuits*, 35(8):1146–1152, 2000. 3

[28] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017. 7

[29] Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. Looking around the corner using transient imaging. In *2009 IEEE 12th International Conference on Computer Vision*, pages 159–166. IEEE, 2009. 3

[30] 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, nov 2013. 3

[31] Martin Laurenzis. Single photon range, intensity and photon flux imaging with kilohertz frame rate and high dynamic range. *Optics Express*, 27(26):38391, dec 2019. 3

[32] Timothy Lee. Most lidars today have between 1 and 128 lasersthis one has 11,000. *Ars Technica*, Jan 2020. Accessed Nov 16, 2020. 8

[33] David B Lindell and Gordon Wetzstein. Three-dimensional imaging through scattering media based on confocal diffuse tomography. *Nature communications*, 11(1):1–8, 2020. 1

[34] Scott Lindner, Chao Zhang, Ivan Michel Antolovic, Martin Wolf, and Edoardo Charbon. A 252 x 144 spad pixel flash lidar with 1728 dual-clock 48.8 ps tdcs, integrated histogramming and 14.9-to-1 compression in 180nm cmos technology. In *2018 IEEE Symposium on VLSI Circuits*, pages 69–70. IEEE, 2018. 1

[35] Xiaochun Liu, Ibón Guillén, Marco La Manna, Ji Hyun Nam, Syed Azer Reza, Toan Huu Le, Adrian Jarabo, Diego Gutierrez, and Andreas Velten. Non-line-of-sight imaging using phasor-field virtual wave optics. *Nature*, 572(7771):620–623, 2019. 1

[36] Sizhuo Ma, Shantanu Gupta, Arin C Ulku, Claudio Bruschini, Edoardo Charbon, and Mohit Gupta. Quanta burst photography. *ACM Transactions on Graphics (TOG)*, 39(4):79–1, 2020. 2, 3

[37] Raffi Mardirosian. Why Apple chose digital lidar. *Ouster Blog*, Apr 2020. Accessed Nov 16, 2020. 8

[38] Ben Mildenhall, Jonathan T Barron, Jiawen Chen, Dillon Sharlet, Ren Ng, and Robert Carroll. Burst denoising with kernel prediction networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 8, 7

[39] Kazuhiro Morimoto, Andrei Ardelean, Ming-Lo Wu, Arin Can Ulku, Ivan Michel Antolovic, Claudio Bruschini, and Edoardo Charbon. Megapixel time-gated SPAD image sensor for 2d and 3d imaging applications. *Optica*, 7(4):346, apr 2020. 3, 16, 17

[40] S.K. Nayar and T. Mitsunaga. High dynamic range imaging: spatially varying pixel exposures. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*. IEEE Comput. Soc, 2000. 3

[41] Matthew O'Toole, Felix Heide, David B Lindell, Kai Zang, Steven Diamond, and Gordon Wetzstein. Reconstructing transient images from single-photon sensors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1539–1547, 2017. 3

[42] Sylvain Paris. A gentle introduction to bilateral filtering and its applications. In *ACM SIGGRAPH 2007 courses*, pages 3–es. 2007. 7

[43] Matteo Perenzoni, Nicola Massari, Daniele Perenzoni, Leonardo Gasparini, and David Stoppa. A 160 x 120 pixel analog-counting single-photon imager with time-gating and self-referenced column-parallel a/d conversion for fluorescence lifetime imaging. *IEEE Journal of Solid-State Circuits*, 51(1):155–167, 2015. 2

[44] Davide Portaluppi, Enrico Conca, and Federica Villa. 32 × 32 CMOS SPAD imager for gated imaging, photon timing, and photon coincidence. *IEEE Journal of Selected Topics in Quantum Electronics*, 24(2):1–6, mar 2018. 16

[45] Joshua Rapp, Julian Tachella, Yoann Altmann, Stephen McLaughlin, and Vivek K Goyal. Advances in single-photon lidar for autonomous vehicles: Working principles, challenges, and recent advances. *IEEE Signal Processing Magazine*, 37(4):62–71, 2020. 1

[46] Julián Tachella, Yoann Altmann, Nicolas Mellado, Aongus McCarthy, Rachael Tobin, Gerald S Buller, Jean-Yves Tourneret, and Stephen McLaughlin. Real-time 3D reconstruction from single-photon lidar data using plug-and-play point cloud denoisers. *Nature communications*, 10(1):1–6, 2019. 1

[47] The ImageMagick Development Team. Imagemagick. 7

[48] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, Lei Zhang, Bee Lim, et al. Ntire 2017 challenge on single image super-resolution: Methods and results. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017. 7

[49] Alex Turpin, Gabriella Musarra, Valentin Kapitany, Francesco Tonolini, Ashley Lyons, Ilya Starshynov, Federica Villa, Enrico Conca, Francesco Fioranelli, Roderick Murray-Smith, et al. Spatial images from temporal data. *Optica*, 7(8):900–905, 2020. 1

[50] Arin Can Ulku, Claudio Bruschini, Ivan Michel Antolovic, Yung Kuo, Rinat Ankri, Shimon Weiss, Xavier Michalet, and Edoardo Charbon. A 512x512 spad image sensor with integrated gating for widefield flim. *IEEE Journal of Selected Topics in Quantum Electronics*, 25(1):1–12, jan 2019. 1

[51] Feng Yang, Yue M Lu, Luciano Sbaiz, and Martin Vetterli. Bits from photons: Oversampled image acquisition using binary poisson statistics. *IEEE Transactions on image processing*, 21(4):1421–1436, 2011. 5

[52] Junko Yoshida. Breaking Down iPad Pro 11's LiDAR Scanner. *EE Times*, Jun 2020. Accessed Nov 16, 2020. 1, 8

[53] Majid Zarghami, Leonardo Gasparini, Matteo Perenzoni, and Lucio Pancheri. High dynamic range imaging with TDC-based CMOS SPAD arrays. *Instruments*, 3(3):38, aug 2019. 3

[54] Hang Zhao, Boxin Shi, Christy Fernandez-Cull, Sai-Kit Yeung, and Ramesh Raskar. Unbounded high dynamic range photography using a modulo camera. In *2015 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10. IEEE, 2015. 3