

Photon Inhibition for Energy-Efficient Single-Photon Imaging

Lucas J. Koerner¹, Shantanu Gupta², Atul Ingle³, and Mohit Gupta²

¹ Department of Electrical and Computer Engineering, University of St. Thomas, St. Paul, MN, 55105, USA koerner.lucas@stthomas.edu

² Department of Computer Sciences, University of Wisconsin-Madison, Madison, WI, 53706, USA {sgupta,mohitg}@cs.wisc.edu

³ Department of Computer Science, Portland State University, Portland, OR, 97201, USA ingle2@pdx.edu

Abstract. Single-photon cameras (SPCs) are emerging as sensors of choice for various challenging imaging applications. One class of SPCs based on the single-photon avalanche diode (SPAD) detects individual photons using an avalanche process; the raw photon data can then be processed to extract scene information under extremely low light, high dynamic range, and rapid motion. Yet, single-photon sensitivity in SPADs comes at a cost — each photon detection consumes more energy than that of a CMOS camera. This avalanche power significantly limits sensor resolution and could restrict widespread adoption of SPAD-based SPCs. We propose a computational-imaging approach called *photon inhibition* to address this challenge. Photon inhibition strategically allocates detections in space and time based on downstream inference task goals and resource constraints. We develop lightweight, on-sensor computational inhibition policies that use past photon data to disable SPAD pixels in real-time, to select the most informative future photons. As case studies, we design policies tailored for image reconstruction and edge detection, and demonstrate, both via simulations and real SPC captured data, considerable reduction in photon detections (over 90% of photons) while maintaining task performance metrics. Our work raises the question of “which photons should be detected?”, and paves the way for future energy-efficient single-photon imaging. Source code for our experiments is available at <https://visionlab.com/project/inhibition>.

Keywords: single-photon imaging · SPADs · energy-efficient vision

1 Introduction

From vacuum tube-based single-photon detectors invented in the early 20th century [25], to 3D depth sensing LiDAR cameras found in today’s smartphones [28], single-photon camera (SPC) technology has come a long way in terms of pixel resolution and commercial availability for a variety of applications. Thanks to CMOS-compatible single-photon avalanche diode (SPAD) arrays, SPCs are being increasingly used not only in niche fields such as scientific imaging and biomedical

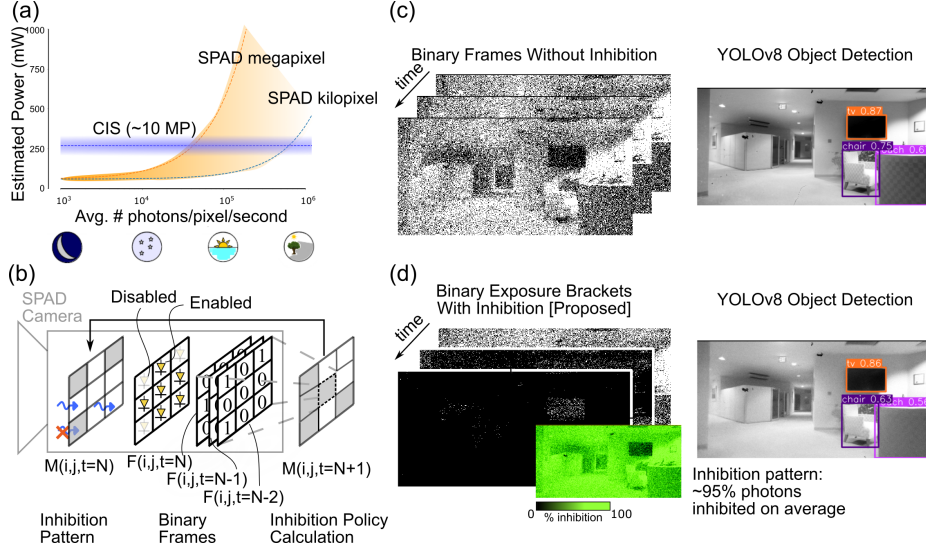


Fig. 1: Photon inhibition for resource-efficient passive SPAD imaging. (a) Unlike conventional CMOS cameras (CIS), the energy consumption in SPAD cameras increases with scene brightness, severely limiting the applicability of high resolution SPAD cameras in resource-constrained applications. (b) We expand the conventional imaging pipeline to incorporate “inhibition” that electronically enables or disables individual pixels to limit bandwidth and power consumption. Our method relies on lightweight mathematical operations called “inhibition policies” that update the inhibition patterns based on the history of photon detections. Inhibition policies can be optimized for image SNR or for downstream vision tasks. (c,d) Object detection, a high-level vision task, is successful even with a large fraction of photons inhibited.

microscopy [49], but more widely for consumer photography [37]. SPAD-based SPCs have recently been fabricated into kilo-to-megapixel format arrays that are now commercially available [28, 37]. The extreme sensitivity and high speed can benefit passive low-light computer vision tasks, particularly in the presence of rapid scene or camera motion [35], enable wide dynamic-range imaging [14, 31] and photon-starved active imaging applications such as 3D imaging (LiDAR) [21] and fluorescence microscopy [49].

SPAD camera pixels detect individual photons with extremely high frame rates by exploiting avalanche multiplication. On one hand, being able to detect individual photons opens up new possibilities and capabilities for computer vision systems. On the other hand, this also presents a unique challenge: Every photon-induced avalanche comes with a non-negligible energy cost, which is a challenge exclusive to SPAD-based cameras. This flux-dependent photon detection power is a significant fraction of total power consumption in today’s SPAD cameras and impedes further increases in their spatial resolution [24, 37, 41, 42, 47]. For example, extrapolating the avalanche power of a recent SPAD sensor [47] to ~ 10 ’s of megapixels format predicts a power consumption of several watts in

bright light⁴ (as illustrated in Fig. 1(a)), far exceeding the power consumption of modern CMOS image sensors (CIS) of around 300 mW [45].

We address this problem by being selective about which photons are detected on the SPAD sensor while maintaining good performance on various computer vision tasks. To reduce avalanche power consumption and decouple it from photon flux, we propose a computational imaging technique called *photon inhibition* where individual SPAD pixels are disabled adaptively based on previous photon detections in their spatio-temporal neighborhoods. Electrically disabling pixels prevents photon detections to inhibit any avalanche power or subsequent processing. We implement lightweight on-sensor computations called *inhibition policies* (Fig. 1(b)) to determine, in real-time and at single-pixel and single-frame granularity, which SPAD pixels to enable or disable. Our method is inspired by retinal pre-processing of the human visual system where retinal neurons aggregate photon information over small spatio-temporal neighborhoods to cause neighboring retinal cells to become less sensitive to incident light [12, 16].⁵

Given that SPADs introduce a new challenge of flux-dependent power consumption, we establish, from first principles, novel *energy-aware* imaging performance metrics for *resource-constrained* single-photon imaging. Based on these metrics, we design families of inhibition policies that distribute photon detections in space and time based on imaging / vision task goals and energy consumption constraints. A critical consideration in the design of inhibition policies stems from the observation that these policies are meant to control (enable / disable) the sensor (Fig. 1(b)). Therefore, it is *essential* for these policies to be extremely lightweight since these need to be implemented on sensor with very limited compute and memory resources. Furthermore, inhibition policies must execute at ultra-low latency to keep up with high-speed photon detections (reaching up to 100 kHz.). Fortunately, since the raw data output from a SPAD-based SPC consists of binary-valued image frames, SPCs are naturally suited to real-time calculations at the image sensor plane under tight compute and memory budgets. The proposed inhibition policies are lightweight, requiring only simple arithmetic and Boolean operations computed over local spatio-temporal neighborhoods, and thus amenable to in-pixel implementation [2, 7].

In simulations and real experimental data, we show that our inhibition policies allocate photon detections to sensor pixels in a way that reduces detection energy for a given accuracy level of various vision tasks. Our results show consequential energy savings when compared to a conventional capture scheme for tasks of (i) image reconstruction: 42% fewer photon detections at equal image quality; (ii) edge detection: an edge sensitive inhibition policy reduces detections by 30% at equal F-score; and (iii) YOLOv8 object detection: remains success-

⁴ It has been shown, perhaps counter-intuitively, that SPADs do not saturate even under extremely bright conditions [26, 27]. Therefore, SPADs are not restricted to low-flux environments, but are being considered for vision applications across a wide dynamic range of lighting conditions (e.g., from a dark tunnel to bright sunlight).

⁵ We borrow the term “inhibition” from the phenomenon of “lateral inhibition” found in biological vision systems [3].

ful with 95% of photons inhibited under camera motion of a real-world SPC (Fig. 1(c,d)). Through experiments with photon streams captured using real SPAD camera hardware over a wide range of illumination conditions, we show that our proposed inhibition policies preserve low-light details and, in bright-light, decouple flux and detection energy.

Scope and Limitations: There are several competing image sensor technologies today that resolve single-photons while capturing binary-valued raw frames at rates exceeding thousands of frames/second. SPCs based on “jots” [32] that do not rely on avalanche multiplication do not suffer from flux-dependent power consumption as SPADs. Jots are an exciting technology, especially in scenarios that require high resolution and high dynamic range imaging [33]. In this paper, we focus on SPADs, due to their rapidly rising availability and commercial interest [6, 28], and benefits over conventional CMOS sensors, both in low-light and bright scenes for a variety of computer vision tasks [10, 35].

Our goals in this paper are to (a) raise the question of “which photons should be detected?” in the context of energy-efficient single-photon imaging, (b) establish a design space and metrics to evaluate various inhibition policies, and (c) propose *plausible* policies that respect practical hardware limitations for future on-chip implementation. We emphasize that the inhibition policies proposed in this paper are not necessarily optimal. This work is just a first step towards demonstrating that it is possible to achieve high performance with SPADs, while maintaining low power consumption via photon inhibition.

2 Related Work

Hardware approaches for reduced energy consumption: There is a strong dependence of SPAD power consumption on the pixel size — the smaller the pixel, the lower the avalanche energy [39]. Although recent developments in SPAD pixel technology have reduced pixel sizes to below $4\mu\text{m}$ [40], avalanche energy still contributes a large fraction of the total power consumption in a SPAD sensor [44]. SPAD design optimizations have reduced the charge per avalanche by RF modulation of the bias voltage [52], minimization of the junction capacitance [42], and smart [5] and fast [54] quenching circuits. Circuit architectures may require spatial and/or temporal co-incidence [22] to reduce energy downstream in the processing chain, but avalanche energy remains. Our work complements existing hardware approaches by preventing avalanches altogether to reduce illumination-dependent energy consumption.

SPAD dead-time and clocked recharge: SPADs require recharge after an avalanche-inducing photon detection during which subsequent photons are not recorded. This dead-time inhibits photons at high exposure [24, 27], yet, power consumption remains excessive when the average inter-photon arrival interval is shorter than the SPAD dead time [38, 42]. Power consumption at high photon flux can be reduced by controlling the global rate of SPAD recharge (“clocked recharge”) [38, 47]. Within a single recharge period a SPAD detects at most a single photon; subsequent photon arrivals do not induce an avalanche, thus power is

reduced. To further limit avalanches, clocked recharge has been combined with a limit on the number of detections and time to saturation circuitry to measure the intensity of the saturated pixels [41], coarse pixel-wise exposure control [42], and a sequence of recharge periods similar to exposure bracketing [47]. These methods could be considered special cases of inhibition policies that are global and do not adapt to scene content, and therefore, are limited in flexibility to trade power and measurement fidelity. Consequently, in these methods, the power of avalanches remains a considerable fraction of total SPAD sensor power consumption [47]. For example, in this work we show considerable power savings, reaching up to 90%, for an object detection task.

Resource-aware imaging: Event-based vision sensors reduce power consumption by only transmitting scene information when an intensity change is detected [17]. This idea has recently been applied to SPAD arrays to reduce power consumption due to data transfer [11, 46]. In contrast, our method reduces power due to detection by selectively disabling photodiodes based on photon history over small spatial and temporal neighborhoods. Miniaturized cameras with constraints on compute energy have transferred processing to the optical domain [30]. While we focus on passive imaging, depth sensing with SPADs and an active pulsed light source has related constraints such as acquisition time and laser power. Adaptive gating reduces acquisition time [43] and optimal allocation of the laser dwell time among pixels improves data quality [36, 48].

3 Observation Model

During an exposure time T , a photon flux of ϕ results in an average number of photon conversions, or exposure, of $H = \phi T$ (we fold the sensor’s *photon detection probability*, or *PDP*, into the definition of ϕ , meaning it represents an effective flux). The distribution of photon conversions, K , follows the Poisson distribution given by $P(K = k; H) = \frac{H^k e^{-H}}{k!}$. During each binary exposure period, a SPAD pixel records a ‘1’ if at least one photon was incident during that period, and ‘0’ otherwise. The probability of detecting at least one photon is given by $Y := 1 - P(K = 0; H) = 1 - e^{-H}$. Hence, the SPAD pixel readout in each binary frame is a Bernoulli random variable $B \sim \text{Bernoulli}(Y)$. Multiple exposure time windows, or measurements (W), are recorded to reduce noise with the total number of detections

$$D := \left[\sum_{n=1}^W B_n \right] \sim \text{Binomial}(W, 1 - e^{-H}). \quad (1)$$

We estimate the probability of detection and flux from a measurement D as

$$\hat{Y} = \frac{D}{W} \quad (2)$$

$$\hat{H} = -\ln(1 - \hat{Y}). \quad (3)$$

Changes under inhibition: Inhibition is represented by a binary state variable M at every pixel, with $M_n = 1$ denoting enabled for the n -th measurement period. M is a random variable when an adaptive or data-dependent inhibition policy is used. The total number of measurements changes to $W_{inh.} := \sum_{n=1}^W M_n \leq W$. The number of detections becomes

$$D_{inh.} := \left[\sum_{n=1}^W M_n B_n \right] = \left[\sum_{\substack{n=1 \\ M_n=1}}^W B_n \right] \sim \text{Binomial}(W_{inh.}, 1 - e^{-H}), \quad (4)$$

and the flux is estimated similar to Eqs. (2) and (3). The second summation conveys that, with inhibition, the measurements when the pixel is enabled match the original model in Eq. (1) – a result of the memoryless property of the Poisson arrival process. The model above requires that transitions in M are synchronized with the clock signal used to gate the exposure, so that the PDP is not changed by inhibition. A second assumption is that the time for SPAD recharge is small relative to the clock period. This is a desired property for passive SPAD-based imaging, and holds for many state-of-the-art SPADs [42, 44, 47]. It implies that PDP is approximately constant in time and does not depend on prior pixel state.

4 Energy-Aware Performance Metrics

The exposure-referred signal-to-noise ratio (SNR_H) is commonly used to evaluate single-photon sensor performance, and can be computed as the ratio of the true exposure H and the root-mean-squared error in the estimated exposure $\sqrt{\mathbb{E}[(\hat{H} - H)^2]}$ [15, 55]:

$$\text{SNR}_H = \frac{H}{\sqrt{\mathbb{E}[(\hat{H} - H)^2]}} = H \sqrt{\frac{W}{e^H - 1}}. \quad (5)$$

At low incident flux, SNR_H is low due to shot noise. SNR_H improves as the likelihood of a photon detection increases until, in bright light with $H > 1.6$, the SNR_H degrades due to “soft” saturation of the response [27, 34].

We propose two new *energy-aware* modifications to SNR_H to incorporate SPAD energy costs. First, we propose a detection efficiency metric $\text{SNR}_{H/D}^2$ defined as the square of SNR normalized by the expected number of detections $\mathbb{E}[D] = W(1 - e^{-H})$:

$$\text{SNR}_{H/D}^2 := \frac{\text{SNR}_H^2}{\mathbb{E}[D]} = \frac{H^2 e^{-H}}{(1 - e^{-H})^2}. \quad (6)$$

Fig. 3(a) shows the SNR_H (black) and the detection efficiency (red) versus the average photon arrivals per period (H). When $H \ll 1$ all detections contribute significant information and SNR_H^2 increases linearly, similar to an ideal non-saturating sensor only limited by Poisson noise. Accordingly, $\text{SNR}_{H/D}^2 \approx 1$, the

upper bound of this metric. At larger exposure values, beginning around $H \approx 0.5$, the sensor begins to saturate which slows the growth of SNR_H and thus degrades the detection efficiency.

A separate constraint is the total number of recharge periods during which a pixel is enabled and can measure either ‘0’ or ‘1’. This number of measurements (W) may be limited due to the energy to read out a frame, the depth of an in-pixel counter, and/or the maximum allowable sensing latency due to motion blur. We establish a second metric, *measurement efficiency* defined as the square of SNR normalized by the number of measurement windows:

$$\text{SNR}_{H/W}^2 := \frac{\text{SNR}_H^2}{W} = \frac{H^2 e^{-H}}{1 - e^{-H}}. \quad (7)$$

Fig. 3(a) (blue) shows sub-optimal measurement efficiency at both low and high exposures with the best efficiency at $H = 1.59$, $Y = 0.80$, as demonstrated in [8].

In an oracle setting with a known image, one can analytically derive non-uniform allocations of measurements W_i across the pixels i , constraining the total expected detections to a fixed value and optimizing SNR_H or mean-squared-error (see the supplement). The loss metric may also be defined relative to the binary rate Y instead of H as above. In that case a useful base metric could be *entropy* instead of SNR_H [19].

5 Spatio-temporal Inhibition Policies

We now propose policies that calculate a spatio-temporal inhibition pattern for each pixel and each frame based on the history of photon frames and patterns. Following Sec. 3, we define an inhibition pattern using a binary-valued tensor M , where $M(i, j, t) = 0$ if pixel (i, j) is disabled in the t^{th} frame, and $M(i, j, t) = 1$ otherwise. All pixels are initially enabled, and on-sensor calculations modify M over time. The binary photon cube is defined as $F(i, j, t) = 1$ if the pixel is enabled (*i.e.*, $M(i, j, t) = 1$) and a photon is detected at pixel location (i, j) in the t^{th} frame, otherwise it is zero.

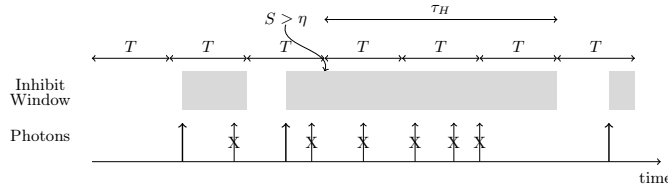


Fig. 2: Calculation-based inhibition with dead time. Arrows represent photons with an ‘X’ for inhibition. T indicates the clocked recharge period. A score, S , calculated from past frames determines if future measurements are enabled or disabled.

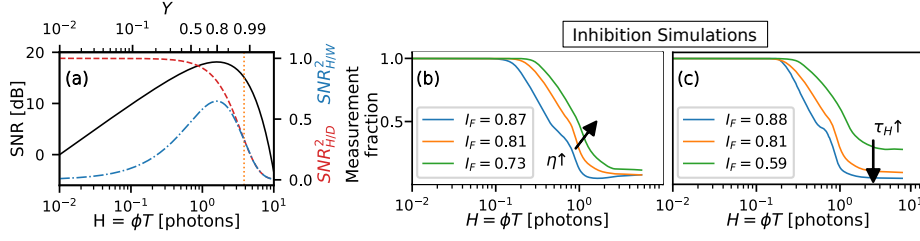


Fig. 3: Efficiency metrics and inhibition policies that track the metrics: (a) The SNR_H in dB (black), the detection efficiency (red, $---$), and the measurement efficiency (blue, $- \cdot -$) versus the exposure with $W = 100$ measurements. The binary rate $Y = 1 - e^{-H}$ is indicated on the top axis. The vertical dotted line indicates the exposure and the binary rate at which the SNR_H degrades by 3 dB from the peak SNR_H . (b,c) Monte Carlo simulations of binary images using calculation-based inhibition policies demonstrate how the allocation of measurements versus the pixel exposure depends upon tuning parameters. (b) the inhibition threshold η adjusts the exposure level at which pixels are inhibited to allow the measurement fraction (the ratio of active measurements to total number of frames) to follow the $\text{SNR}_{H/D}^2$ curve in (a). A smaller threshold more aggressively inhibits photons. (c) demonstrates the impact of the hold-off time, τ_H , on the number of measurements allocated to the brightest pixels. The legend indicates the total fraction of photons inhibited as I_F . (b,c) show smoothed curves (Lowess filter, fraction of 1/5) of the measurement fraction vs. H .

Fig. 1(b) shows the components of a photon inhibition processing layer, including the binary frames, F , and inhibition pattern M . For ease of on-sensor implementation, we focus on policies that operate on small and local spatio-temporal neighborhoods of fixed sizes. We rely on local arithmetic and Boolean computations and comparison operations, consistent with current in-pixel computational capabilities [2, 9, 24]. Fig. 2 shows a proposed on-sensor calculation approach that operates in a streaming fashion as frames accumulate to calculate an inhibition score, S , as the result of a spatio-temporal filter of the binary frames and inhibition pattern. The score at each pixel is calculated as

$$S(i, j, t) = K * [(2F(i, j, t) - 1) \cdot M(i, j, t)] \quad (8)$$

which applies a spatio-temporal filtering kernel, K , of dimensions L, H, T to a ternary representation of the pixel result (1, 0, or -1 for a detection, a disabled pixel, or a measurement that does not detect a photon, respectively). The kernel K can typically be separated into spatial and temporal components as $K = K_s \otimes K_t$ with dimensions $L \times H \times 1$ and $1 \times 1 \times T$, respectively. After each binary frame, the score is compared to a threshold η and the pixel is disabled for the subsequent τ_H frames: $M(i, j, t') = 0$ for $\{t' | t + 1 \leq t' \leq t + 1 + \tau_H\}$ if $S(i, j, t) > \eta$. Observe in Fig. 3(b,c) that decreasing η and increasing τ_H can be used to attain more aggressive inhibition with a larger fraction of photons being inhibited. The binary rate of each pixel is estimated (using Eq. (4)) as the ratio of detections to (active) measurements: $\hat{Y}(i, j, t) = \sum_{t'} F(i, j, t') / \sum_{t'} M(i, j, t')$. This

calculation requires a record of the inhibition history which could be accumulated by an in-pixel counter or recreated in a downstream processor if all binary frames are read out.

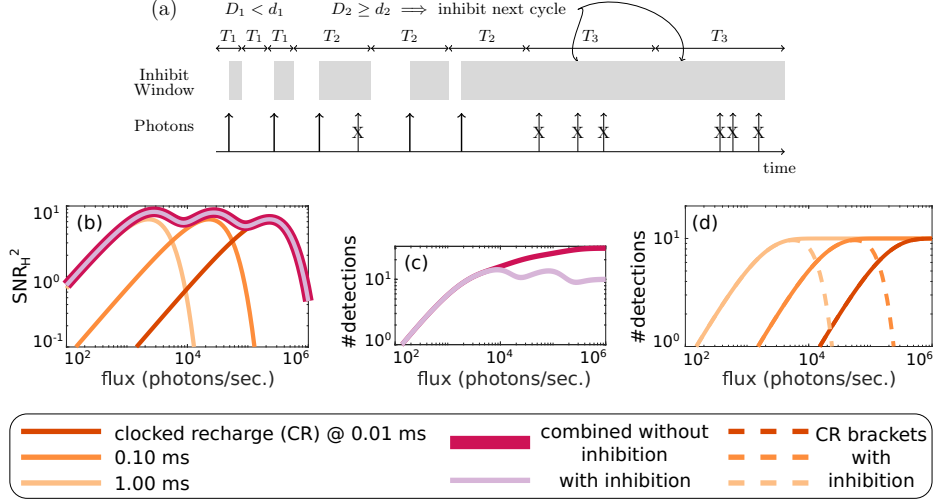


Fig. 4: Saturation look-ahead inhibition. (a) When detections D_i of cycle i exceed a threshold, d_i , the photons arriving during the next recharge cycle, $i+1$, are inhibited. For this drawing $[d_1, d_2] = [3, 3]$. (b, c, d) Example with three recharge period settings and brackets of $W_i = 10$ measurements each. The saturation thresholds are set as $[d_1, d_2] = [7, 7]$. The brackets are combined linearly, weighted by SNR_H^2 [18]. (b, c) With negligible effect on reconstructed SNR relative to conventional clocked-recharge, significantly fewer photon detections occur under the inhibition policy in high-flux settings. (d) The individual brackets are effectively disabled in flux regimes with low detection efficiency, and the detection rate closely tracks the original SNR_H^2 curve. Simulation details are in the supplement.

Example policies: 1) *Single-pixel dead time:* A spatial kernel K_s of dimensions 1×1 may mimic the dead time of a passive SPAD without influence from neighboring pixels. Unlike SPAD recharge generated dead time, the temporal kernel may be extended (e.g., N frames long) to establish a rate threshold for inhibition with reduced quantization noise. 2) *Local spatio-temporal averaging:* Single binary frames are inherently noisy. An inhibition policy that calculates spatio-temporal averages to estimate the local photon rate may reduce the impact of noise on the inhibition pattern. Another benefit, as shown in Fig. 3(b,c), is that the stochastic nature of binary frames smooths the distribution of measurements versus pixel flux. Discontinuities (“dips”) in SNR versus photon flux are undesirable due to the potential for artifacts. 3) *Edge enhancement:* Pixels may be inhibited if a local neighborhood has little spatial variation in photon rate. This can be achieved, for instance, through a spatial filter K_s in Eq. (8) which acts

like a Laplacian filter. Such a strategy may enhance the fidelity of edges in the image while focusing fewer resources on regions with constant illumination.

Fig. 4 presents a second proposed inhibition policy called *saturation look-ahead*. This policy combines exposure brackets and calculation-based inhibition for a light-weight single-pixel inhibition policy. This policy proceeds as a sequence of cycles (indexed by i) of binary frames where each binary frame within each cycle uses the same exposure time. Cycle exposure times T_i progressively increase ($T_1 < T_2 < T_3 \dots$) so that measurements taken in an earlier cycle may predict low detection efficiency (near saturation) at longer exposure times and disable the pixel in these subsequent cycles. The exposure level thresholds for inhibition would typically be set such that the number of photons detected at a given flux level tracks the SNR_H^2 (see Fig. 4(b)), but may be adjusted further based on the relative importance of power consumption, sensing latency, and SNR in an actual application setting.

6 Simulation-based Evaluation

We use Monte Carlo simulations to generate sequences of grayscale binary frames from a dataset of RGB images [1]. The inhibition policies evaluated extend the baseline inhibition generated by clocked recharge. The inhibition score and patterns for various policies and tuning parameters are calculated from these binary frame sequences. Once inhibition patterns are determined, performance is evaluated by tabulating detections, measurements, and image quality or vision task performance for each step in the sequence (see the supplement for details).

Spatio-Temporal Policies for Imaging: Handcrafted spatial kernels (3x3) were combined with an averaging temporal kernel of length 4 to form lightweight inhibition policies that allocate pixel measurements as described in earlier sections for improved image reconstruction. Fig. 5 summarizes the simulation results. Fig. 5(d,e) display reduction in photon detections at equal structural similarity index measure (SSIM) [51], enabled by disabling bright pixels, for an exposure bracketing sequence. Intensity estimates from each bracket are combined using SNR^2 weighting [18] and then converted to a binary rate estimate at the center exposure level of 1 ppp. The proposed policy demonstrates an average reduction in detections of 42% as compared to no inhibition. Fig. 5(f,g) evaluates a single exposure level (1.0 ppp) which is a more challenging scenario, yet the proposed policy still reduces detections by 14% at SSIM=0.7. See supplement S 3.4 for more examples and S 3.5 for simulations of high dynamic range images.

Edge Detection: The BSDS500 dataset with ground truth boundaries [1] was used to study energy-efficient edge detection via photon inhibition. Binary rate images were processed by pre-trained holistically-nested edge detection (HED) [53] with the resulting edge maps compared to ground truth by the structured edge detection toolbox [13]. Fig. 6 shows the optimal image scale (OIS) F-score versus the average detections per pixel. Interpolated curves (not shown) allow for translating along horizontal lines of equal task performance to assess differences in avalanche energy. At low photon counts the proposed edge-enhancing policy

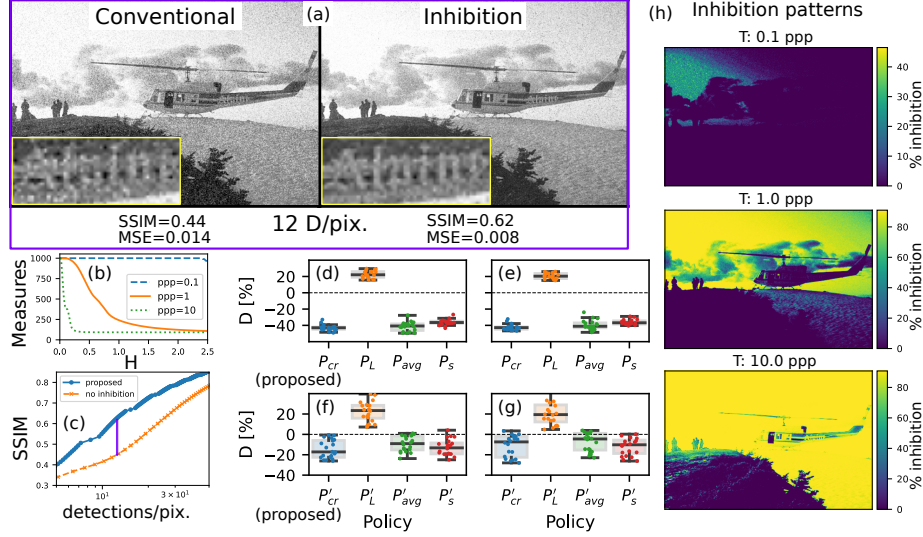


Fig. 5: Power-efficient static imaging via inhibition. (a) Images from an exposure bracketing sequence (average exposure of 0.1, 1.0, 10.0 photons per pixel (ppp)) using clocked recharge without (conventional) and with inhibition. (b) the distribution of measurements based on pixel intensity with H at 1.0 ppp. (c) image quality (SSIM) at equal detections/pixel. (d-g) Handcrafted policies are tested over 20 images and assessed by reductions in detections (D) at equal SSIM. The top row (d),(e) use exposure bracketing; (f),(g) use a single exposure of 1.0 ppp. (d),(f) at SSIM=0.7 and (e),(g) at SSIM=0.8. The box shows quartiles with the center line at the median. The proposed policy, P_{cr} , is a 3×3 spatial kernel that emphasizes the center pixel ($\times 8$) and includes the 8 neighbors ($\times 1$) (see the supplement for policy details). (h) average inhibition patterns for each exposure time. The top most pattern inhibits the brightest pixels only (maximum of $\sim 60\%$ inhibition, primarily in the sky). For the longest exposure time the inhibition pattern allocates measurements primarily to the darkest pixels.

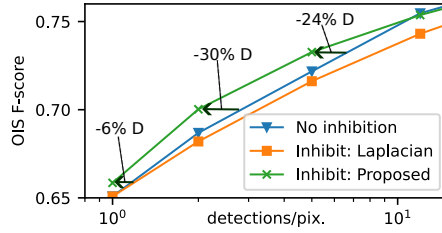


Fig. 6: Energy-efficient edge detection by inhibition: Edge detection F-scores averaged over 19 images versus average detections/pixel. Horizontal arrows show reduction in detections enabled by inhibition at equal task performance. At and beyond 30 D/pix. no inhibition and the proposed policy are nearly equivalent and plateau at 200 D/pix. The edge detector returns $F = 0.813$ on the original images of this set.

demonstrates a 30% reduction in detections on single exposure time captures. This policy extends the on-sensor calculation approach by calculating two scores, the conventional 3×3 Laplacian [50] as S_1 and a 3×3 averaging filter as S_2 . The final inhibition decision is the Boolean operation of these scores as $((\eta_1 < S_1 < \eta_2) \wedge (S_2 > \eta_3)) \vee (S_2 > \eta_4)$. The Laplacian policy alone performs poorly as photons in dim regions with minimal spatial variation are inhibited – disabling dim pixels is energy inefficient. Complete policy descriptions are in the supplement.

7 Experiments on Dynamic Scenes

Many real-world scenes contain significant motion even with the high frame rate of a SPAD camera. Burst reconstruction algorithms yield high-quality images from sequences of binary frames [34,35] – we investigate their compatibility with the inhibited photon detection data. In particular, we focus on the saturation look-ahead policy of Fig. 4(a), *applied independently at every pixel*, and therefore an example of an adaptive single-pixel temporal policy.

To avoid losing salient information under motion, the inhibition policy must limit periods of extended dead time. In the context of the saturation look-ahead policy this limits the total exposure length of the bracketing sequence, since we implicitly assume flux to be constant within each sequence. In our experiments, we use a Fibonacci bracketing [20] sequence $T := \{1, 1, 2, 3, 5, 8, 13, 21\}$ ⁶, denoted in the units of a single minimal exposure time. Every measured bracket is converted to a maximum-likelihood estimate of flux, which is then supplied to the quanta burst photography algorithm [34]. The flux estimator from bracketed measurements is described in detail in supplement Sec. S7.2.

Results with the SwissSPAD2 sensor: The SwissSPAD2 sensor [49] is a prototype SPAD pixel array that can produce binary frames at a rate up to 97,700 FPS, with a resolution of 512×256 . In our experiments, we use binary frames captured directly (without inhibition) by the SPAD array as reference data and emulate on-sensor saturation look-ahead inhibition in software. As a pre-processing step, measurements at hot pixels are replaced with their nearest neighbors.

Fig. 7 shows the results of burst reconstruction under three lighting conditions. The raw data is a sequence of $>580,000$ binary frames with scene radiance increasing rapidly by orders of magnitude (Fig. 7(b)), from <1 lux to $>4,000$ lux, measured separately with a light meter. For each of 47 equally-spaced keyframes, centered windows of 12,000 binary frames are extracted and processed as described above. Results for the full sequence can be found in the supplementary material. A static inhibition policy of regular sub-sampling (dropping 9 out of every 10 frames) is also applied, which yields a fixed 90% reduction in both measurements and photon detections under all lighting conditions. Other sub-sampling factors are discussed in the supplement.

⁶ Sequence $T' := \{1, 1, 1, 3, 3, 3, 8, 8, 25\}$ yielded similar results. No extensive search over the policy space was performed.

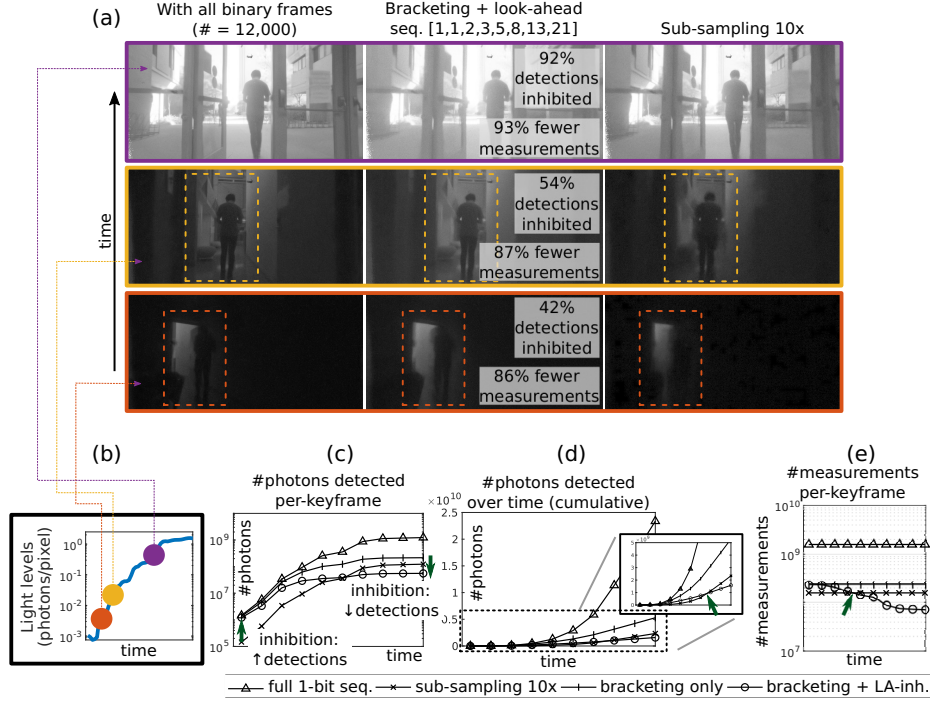


Fig. 7: Adaptive policies on video sequences enable stronger inhibition, preserve low-light details and, in bright-light, decouple flux and detection energy. (a) Burst reconstructions [34] for three keyframes with varying light levels: the top and bottom row differ by ≈ 7 stops. [Images show the detection rates Y and are further gamma-compressed ($\gamma = 0.4$). Complete results included in the supplement.] Left column results are from the original binary frames without inhibition, and the right column after sub-sampling $10\times$ (a fixed 90% inhibition). Middle column represents exposure bracketing combined with saturation look-ahead (see Fig. 4a for description). Under strong light (top row) the results are reasonable with both methods. However, plain sub-sampling loses details in lower light: notice the furniture and a person’s outline in the middle & bottom rows, respectively. Inhibition is instead adaptive to flux. (b) Average exposure level for each keyframe in the sequence. (c,d) Per-keyframe and cumulative detection counts – inhibition ultimately results in fewer photons being detected over the whole sequence. (e) Number of measurements taken for each keyframe; reductions may be translated to energy savings during read-out. Plots in (c,d,e) are sub-sampled for clarity, and crossover points are marked by green arrows.

The top row in Fig. 7(a) shows that under strong light a large fraction of photons ($>90\%$) can be inhibited through the saturation look-ahead policy and still result in good image quality after burst reconstruction, thus spreading photon detections over a longer period of time to reduce avalanche power. Even simple sub-sampling yields good results in bright light, and may work well under controlled illumination conditions. However, the images in the middle and

bottom rows illustrate that this static inhibition policy results in excessive signal loss in lower light leading to a loss of details. Sub-sampling may be enhanced by adjusting exposure time and/or the sub-sampling factor in response to global flux but cannot simultaneously optimize for different light levels within a single frame: over-exposed regions may have clipping artifacts and under-exposed regions are prone to motion blur from incorrect burst reconstruction. The saturation look-ahead policy does adapt to local flux, and allocates relatively more detections to dim regions (see Fig. 1(d)). It thus decouples detection energy from flux (Fig. 7(c); cf. Fig. 1(a)), and results in cumulatively fewer detections than sub-sampling (Fig. 7(d)) due to it being more aggressive in strong light.

Apart from photography, the burst-reconstructed images can be used in computer vision applications, where even stronger inhibition is possible depending on the noise- or blur-robustness of the vision algorithm. Fig. 1(d) shows successful object detection with the YOLOv8 algorithm [29] applied to a burst-reconstructed image, with approximately 95% photon detections inhibited.

8 Limitations and Future Outlook

Implementation costs. Our focus in this paper is on reducing energy consumption due to avalanches. Although avalanches contribute significant energy as compared to on-sensor computations [2, 42, 47], an important next step is to design a holistic model that includes the energy consumption of computations and readout (leveraging Eq. 7). Our in-pixel computations — power-of-two multiplications which simplify to bit shifts, small spatio-temporal kernels no larger than $3 \times 3 \times 4$ — are designed to be lightweight. Fortunately, computational SPAD imagers [2] with in-pixel memory and compute have recently been proposed, with 4×4 block of pixels having a 32-bit CPU and over 10 kbits of memory. See the supplement for an estimate of the required circuitry.

More complete models. The noise model used for efficiency metrics only accounts for photon Poisson noise and the quantization noise of Bernoulli samples. Expanding the noise model to include uncorrected pixel sensitivity variations, crosstalk, and afterpulsing may improve performance [4] by allocating measurements with an awareness of practical sensor limitations. Inhibition modulates the number of Bernoulli trials. An unbiased estimator for data-dependent stopping of Bernoulli trials is known [23], yet is not applied in this work since it does not precisely match our situation. Our simulations (SSIM) suggest that bias is less significant than image noise, yet future analytical studies are needed over a range of conditions.

Generalization to other tasks. To generalize the proposed approaches to a variety of vision tasks, task-specific quality metrics must be defined to pose optimization problems for each task. As an example, for the image reconstruction task, Suppl. Sec. 3.1 shows an analytical optimization using the MSE metric when constrained by photon detections. Other tasks could be approached similarly, yet may need to be optimized empirically.

Acknowledgments

The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources. A.I. was supported in part by NSF ECCS-2138471. S.G and M.G. were supported in part by NSF CAREER award 1943149, NSF award CNS-2107060, and Wisconsin Alumni Research Foundation (WARF).

References

1. Arbeláez, P., Maire, M., Fowlkes, C., Malik, J.: Contour Detection and Hierarchical Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(5), 898–916 (May 2011). <https://doi.org/10.1109/TPAMI.2010.161>
2. Ardelean, A.: Computational Imaging SPAD Cameras. Ph.D. thesis, EPFL (2023). <https://doi.org/10.5075/epfl-thesis-9501>
3. Barlow, H.B.: Summation and inhibition in the frog’s retina. *The Journal of physiology* **119**(1), 69 (1953)
4. Bian, L., Song, H., Peng, L., Chang, X., Yang, X., Horstmeyer, R., Ye, L., Zhu, C., Qin, T., Zheng, D., Zhang, J.: High-resolution single-photon imaging with physics-informed deep learning. *Nature Communications* **14**(1), 5902 (Sep 2023). <https://doi.org/10.1038/s41467-023-41597-9>
5. Boso, G., Buttafava, M., Villa, F., Tosi, A.: Low-cost and compact single-photon counter based on a CMOS SPAD smart pixel. *IEEE Photonics Technology Letters* **27**(23), 2504–2507 (2015)
6. Canon Inc.: Canon Launches MS-500 - The World’s First Ultra-High-Sensitivity Interchangeable-Lens SPAD Sensor Camera. <https://www.usa.canon.com/newsroom/2023/20230801-ms500>, Canon Press Release 8/1/2023. Accessed 2/25/2024.
7. Carey, S.J., Lopich, A., Barr, D.R., Wang, B., Dudek, P.: A 100,000 fps vision sensor with embedded 535 GOPS/W 256×256 SIMD processor array. In: 2013 Symposium on VLSI Circuits. pp. C182–C183 (Jun 2013)
8. Chan, S.H.: What Does a One-Bit Quanta Image Sensor Offer? *IEEE Transactions on Computational Imaging* **8**, 770–783 (2022). <https://doi.org/10.1109/TCI.2022.3202012>
9. Charbon, E., Bruschini, C., Lee, M.J.: 3D-Stacked CMOS SPAD Image Sensors: Technology and Applications. In: 2018 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS). pp. 1–4 (Dec 2018). <https://doi.org/10.1109/ICECS.2018.8617983>
10. Chi, Y., Gnanasambandam, A., Koltun, V., Chan, S.H.: Dynamic low-light imaging with quanta image sensors. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16. pp. 122–138. Springer (2020)
11. Della Rocca, F.M., Mai, H., Hutchings, S.W., Al Abbas, T., Buckbee, K., Tsiamis, A., Lomax, P., Gyongy, I., Dutton, N.A., Henderson, R.K.: A 128×128 SPAD motion-triggered time-of-flight image sensor with in-pixel histogram and column-parallel vision processor. *IEEE Journal of Solid-State Circuits* **55**(7), 1762–1775 (2020). <https://doi.org/10.1109/JSSC.2020.2993722>

12. Diamond, J.S.: Inhibitory interneurons in the retina: Types, circuitry, and function. *Annual Review of Vision Science* **3**(1), 1–24 (2017). <https://doi.org/10.1146/annurev-vision-102016-061345>
13. Dollar, P., Zitnick, C.L.: Structured Forests for Fast Edge Detection. In: *Proceedings of the IEEE International Conference on Computer Vision*. pp. 1841–1848 (2013)
14. Dutton, N.A.W., Al Abbas, T., Gyongy, I., Mattioli Della Rocca, F., Henderson, R.K.: High Dynamic Range Imaging at the Quantum Limit with Single Photon Avalanche Diode-Based Image Sensors. *Sensors* **18**(4), 1166 (Apr 2018). <https://doi.org/10.3390/s18041166>
15. Fossum, E.R.: Modeling the Performance of Single-Bit and Multi-Bit Quanta Image Sensors. *IEEE Journal of the Electron Devices Society* **1**(9), 166–174 (Sep 2013). <https://doi.org/10.1109/JEDS.2013.2284054>
16. Franke, K., Berens, P., Schubert, T., Bethge, M., Euler, T., Baden, T.: Inhibition decorrelates visual feature representations in the inner retina. *Nature* **542**(7642), 439–444 (Feb 2017). <https://doi.org/10.1038/nature21394>
17. Gallego, G., Delbrück, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A.J., Conradt, J., Daniilidis, K., et al.: Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **44**(1), 154–180 (2020)
18. Gnanasambandam, A., Chan, S.H.: HDR Imaging With Quanta Image Sensors: Theoretical Limits and Optimal Reconstruction. *IEEE Transactions on Computational Imaging* **6**, 1571–1585 (2020). <https://doi.org/10.1109/TCI.2020.3041093>
19. Gnanasambandam, A., Chan, S.H.: Exposure-Referred Signal-to-Noise Ratio for Digital Image Sensors. *IEEE Transactions on Computational Imaging* **8**, 561–575 (Jun 2022). <https://doi.org/10.1109/TCI.2022.3187657>
20. Gupta, M., Iso, D., Nayar, S.K.: Fibonacci Exposure Bracketing for High Dynamic Range Imaging. In: *2013 IEEE International Conference on Computer Vision*. pp. 1473–1480. IEEE, Sydney, Australia (Dec 2013). <https://doi.org/10.1109/ICCV.2013.186>
21. Gyongy, I., Dutton, N.A., Henderson, R.K.: Direct time-of-flight single-photon imaging. *IEEE Transactions on Electron Devices* **69**(6), 2794–2805 (2021)
22. Gyongy, I., Martín, G.M., Turpin, A., Ruget, A., Halimi, A., Henderson, R., Leach, J.: High-speed vision with a 3D-stacked SPAD image sensor. In: *Advanced Photon Counting Techniques XV*. vol. 11721, p. 1172105. SPIE (2021)
23. Haldane, J.B.S.: On a Method Of Estimating Frequencies. *Biometrika* **33**(3), 222–225 (Nov 1945). <https://doi.org/10.1093/biomet/33.3.222>
24. Henderson, R.K., Johnston, N., Hutchings, S.W., Gyongy, I., Al Abbas, T., Dutton, N., Tyler, M., Chan, S., Leach, J.: 256×256 40nm/90nm CMOS 3D-stacked 120 dB dynamic-range reconfigurable time-resolved SPAD imager. In: *2019 IEEE International Solid-State Circuits Conference (ISSCC)*. pp. 106–108. IEEE (2019). <https://doi.org/10.1109/ISSCC.2019.8662355>
25. Iams, H., Salzberg, B.: The secondary emission phototube. *Proceedings of the Institute of Radio Engineers* **23**(1), 55–64 (1935)
26. Ingle, A., Sets, T., Buttafava, M., Gupta, S., Tosi, A., Gupta, M., Velten, A.: Passive Inter-Photon Imaging. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 8585–8595 (2021)
27. Ingle, A., Velten, A., Gupta, M.: High Flux Passive Imaging With Single-Photon Sensors. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 6760–6769 (2019)

28. J Yoshida: Breaking Down iPad Pro 11's LiDAR Scanner. <https://www.eetimes.com/breaking-down-ipad-pro-11s-lidar-scanner/>, EE Times 6/5/2020. Accessed 5/6/2021.
29. Jocher, G., Chaurasia, A., Qiu, J.: Ultralytics YOLOv8 (Jan 2023), <https://github.com/ultralytics/ultralytics>, Accessed 3/6/2024
30. Koppal, S.J., Gkioulekas, I., Young, T., Park, H., Crozier, K.B., Barrows, G.L., Zickler, T.: Toward Wide-Angle Microvision Sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35**(12), 2982–2996 (Dec 2013). <https://doi.org/10.1109/TPAMI.2013.22>
31. Liu, Y., Gutierrez-Barragan, F., Ingle, A., Gupta, M., Velten, A.: Single-Photon Camera Guided Extreme Dynamic Range Imaging. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 1575–1585 (2022)
32. Ma, J., Masoodian, S., Starkey, D.A., Fossum, E.R.: Photon-number-resolving megapixel image sensor at room temperature without avalanche gain. *Optica* **4**(12), 1474–1481 (Dec 2017). <https://doi.org/10.1364/OPTICA.4.001474>
33. Ma, J., Zhang, D., Elgendy, O.A., Masoodian, S.: A 0.19e⁻ rms Read Noise 16.7Mpixel Stacked Quanta Image Sensor With 1.1 Mm-Pitch Backside Illuminated Pixels. *IEEE Electron Device Letters* **42**(6), 891–894 (Jun 2021). <https://doi.org/10.1109/LED.2021.3072842>
34. Ma, S., Gupta, S., Ulku, A.C., Bruschini, C., Charbon, E., Gupta, M.: Quanta burst photography. *ACM Transactions on Graphics* **39**(4), 79:1–79:16 (Aug 2020). <https://doi.org/10.1145/3386569.3392470>
35. Ma, S., Mos, P., Charbon, E., Gupta, M.: Burst Vision Using Single-Photon Cameras. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 5375–5385 (2023)
36. Medin, S.C., Murray-Bruce, J., Castañón, D., Goyal, V.K.: Beyond Binomial and Negative Binomial: Adaptation in Bernoulli Parameter Estimation. *IEEE Transactions on Computational Imaging* **5**(4), 570–584 (Dec 2019). <https://doi.org/10.1109/TCI.2019.2913108>
37. Morimoto, K., Iwata, J., Shinohara, M., Sekine, H., Abdelghafar, A., Tsuchiya, H., Kuroda, Y., Tojima, K., Endo, W., Maehashi, Y., Ota, Y., Sasago, T., Maekawa, S., Hikosaka, S., Kanou, T., Kato, A., Tezuka, T., Yoshizaki, S., Ogawa, T., Uehira, K., Ehara, A., Inui, F., Matsuno, Y., Sakurai, K., Ichikawa, T.: 3.2 Megapixel 3D-Stacked Charge Focusing SPAD for Low-Light Imaging and Depth Sensing. In: *2021 IEEE International Electron Devices Meeting (IEDM)*. pp. 20.2.1–20.2.4 (Dec 2021). <https://doi.org/10.1109/IEDM19574.2021.9720605>
38. Morimoto, K., Ardelean, A., Wu, M.L., Ulku, A.C., Antolovic, I.M., Bruschini, C., Charbon, E.: Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications. *Optica* **7**(4), 346–354 (2020). <https://doi.org/10.1364/OPTICA.386574>
39. Morimoto, K., Charbon, E.: A scaling law for SPAD pixel miniaturization. *Sensors* **21**(10), 3447 (2021)
40. Ogi, J., Sano, F., Nakata, T., Kubo, Y., Onishi, W., Koswaththaghe, C., Mochizuki, T., Tashiro, Y., Hizu, K., Takatsuka, T., et al.: A 3.06 μm spad pixel with embedded metal contact and power grid on deep trench pixel isolation for high-resolution photon-counting. In: *2023 International Image Sensor Workshop (IISW)* (May 2023)
41. Ogi, J., Takatsuka, T., Hizu, K., Inaoka, Y., Zhu, H., Tochigi, Y., Tashiro, Y., Sano, F., Murakawa, Y., Nakamura, M., Oike, Y.: A 250fps 124dB Dynamic-Range

- SPAD Image Sensor Stacked with Pixel-Parallel Photon Counter Employing Sub-Frame Extrapolating Architecture for Motion Artifact Suppression. In: 2021 IEEE International Solid- State Circuits Conference (ISSCC). vol. 64, pp. 113–115 (Feb 2021). <https://doi.org/10.1109/ISSCC42613.2021.9365977>
42. Ota, Y., Morimoto, K., Sasago, T., Shinohara, M., Kuroda, Y., Endo, W., Maehashi, Y., Maekawa, S., Tsuchiya, H., Abdelahafar, A., Hikosaka, S., Motoyama, M., Tojima, K., Uehira, K., Iwata, J., Inui, F., Matsuno, Y., Sakurai, K., Ichikawa, T.: A 0.37W 143dB-Dynamic-Range 1Mpixel Backside-Illuminated Charge-Focusing SPAD Image Sensor with Pixel-Wise Exposure Control and Adaptive Clocked Recharging. In: 2022 IEEE International Solid- State Circuits Conference (ISSCC). vol. 65, pp. 94–96 (Feb 2022). <https://doi.org/10.1109/ISSCC42614.2022.9731644>
 43. Po, R., Pediredla, A., Gkioulekas, I.: Adaptive gating for single-photon 3d imaging. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 16354–16363 (2022)
 44. Severini, F., Cusini, I., Berretta, D., Pasquinelli, K., Incoronato, A., Villa, F.: SPAD pixel with sub-ns dead-time for high-count rate applications. *IEEE Journal of Selected Topics in Quantum Electronics* **28**(2: Optical Detectors), 1–8 (2021)
 45. Shenzhen CM Technology company Ltd: 16 MP MIPI Camera Module with SONY IMX206 sensor. <http://www.camera-module.com/product/mipicameramodule/16mp-mipi-camera-module-sony-imx206-sensor.html> (Nov 2023)
 46. Sundar, V., Ardelean, A., Swedish, T., Bruschini, C., Charbon, E., Gupta, M.: Sodacam: Software-defined cameras via single-photon imaging. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV) (2023)
 47. Takatsuka, T., Ogi, J., Ikeda, Y., Hizu, K., Inaoka, Y., Sakama, S., Watanabe, I., Ishikawa, T., Shimada, S., Suzuki, J., Maeda, H., Toshima, K., Nonaka, Y., Yamamura, A., Ozawa, H., Koga, F., Oike, Y.: A 3.36 μm -pitch SPAD photon-counting image sensor using clustered multi-cycle clocked recharging technique with intermediate most-significant-bit readout. In: 2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits). pp. 1–2 (Jun 2023). <https://doi.org/10.23919/VLSITechnologyandCir57934.2023.10185241>
 48. Tilmon, B., Sun, Z., Koppal, S.J., Wu, Y., Evangelidis, G., Zahreddine, R., Krishnan, G., Ma, S., Wang, J.: Energy-Efficient Adaptive 3D Sensing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5054–5063 (2023)
 49. Ulku, A.C., Bruschini, C., Antolovic, I.M., Kuo, Y., Ankri, R., Weiss, S., Michalet, X., Charbon, E.: A 512×512 SPAD Image Sensor With Integrated Gating for Widefield FLIM. *IEEE Journal of Selected Topics in Quantum Electronics* **25**(1), 1–12 (Jan 2019). <https://doi.org/10.1109/jstqe.2018.2867439>
 50. Wang, X.: Laplacian operator-based edge detectors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(5), 886–890 (2007)
 51. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**(4), 600–612 (2004)
 52. Wayne, M.A., Bienfang, J.C., Migdall, A.L.: Low-noise photon counting above 100×106 counts per second with a high-efficiency reach-through single-photon avalanche diode system. *Applied Physics Letters* **118**(13) (2021)
 53. Xie, S., Tu, Z.: Holistically-nested edge detection. In: Proceedings of IEEE International Conference on Computer Vision. pp. 1395–1403 (2015)

- 54. Xu, Y., Lu, J., Wu, Z.: A compact high-speed active quenching and recharging circuit for SPAD detectors. *IEEE Photonics Journal* **12**(5), 1–8 (2020)
- 55. Yang, F., Lu, Y.M., Sbaiz, L., Vetterli, M.: Bits from photons: Oversampled image acquisition using binary poisson statistics. *IEEE Transactions on image processing* **21**(4), 1421–1436 (2011)