

# Parametric Poisson Process Imaging

Donggeek Shin, Ahmed Kirmani, Andrea Colaço, and Vivek K Goyal

Research Laboratory of Electronics  
Massachusetts Institute of Technology

**Abstract**—In conventional 3D imaging, a large number of detected photons is required at each pixel to mitigate the effect of signal-dependent Poisson or shot noise. Parametric Poisson process imaging (PPPI) is a new framework that enables scene depth acquisition with very few detected photons despite significant contribution from background light. Our proposed computational imager is based on accurate physical modeling of the photon detection process using time-inhomogeneous Poisson processes combined with regularization that promotes piecewise smoothness. Simulations demonstrate accurate imaging with only 1 detected photon per pixel.

**Index Terms**—computational 3D imaging, convex optimization, Poisson processes, single-photon avalanche diodes, time-of-flight cameras.

## I. INTRODUCTION

This paper introduces a new conceptual framework for active 3D optical imaging in low-light scenarios. We consider the use of raster-scanned illumination and detection of light with one single-photon avalanche diode (SPAD); the framework applies identically with flood illumination and a SPAD array. Due to the quantum nature of light, the SPAD measures a Poisson process at each spatial location. The framework is called *parametric Poisson process imaging* (PPPI) because the essential idea is that the rate of each Poisson process is a parametric function with parameters that encode features such as scene reflectance, distance to scene points, and background light level. The parametric nature of these functions along with correlations over the spatial dimensions enable effective estimation with only a few detected photons and with significant contribution from background light.

### A. Main contributions

*Conceptual:* We introduce a physically-accurate model for the signal produced by a SPAD under weak lighting that incorporates arbitrary illumination pulse shape, inhomogeneous Poisson process characteristics (shot noise from the quantum nature of light), and detector reset time.

*Algorithmic:* We provide a method for spatiotemporally-regularized estimation of the depth map. Our algorithm produces accurate 3D images from 1 photon per pixel (ppp) data.

*Experimental:* In work detailed elsewhere [1], we achieve very high photon efficiency for 3D depth imaging using raster-scanned illumination with a pulsed laser and detection of the

backscattered light with a single SPAD. Depth estimation at 1 ppp is comparable to a traditional technique with 100 ppp or 17 times narrower pulse width. We also show how the proposed computational imager outperforms state-of-the-art image denoising algorithms.

### B. Related Work

*Algorithmic:* Image acquisition at low-light levels is challenging due to the signal dependent nature of the noise. Denoising of low-light images captured by SPAD sensors is typically accomplished using the Poisson or shot noise model along with sparsity-promoting regularization [2]. These methods assume that the number of collected photons is Poisson distributed with parameter proportional to the true intensity. For depth imaging using a SPAD sensor, first a large number of individual photon arrival times are recorded at every pixel to generate a photon count histogram for each pixel [3], [4]. Since the photon count is very high, the maximum likelihood (ML) depth estimate is obtained by simply applying the matched filter, defined by the pulse waveform, to the count histogram [5]. When the photon count is low, depth estimation will have nonnegligible variance depending on the pulse width. After forming the depth image, spatial averaging is used for further denoising to improve image quality while sacrificing spatial resolution. Thus, forming a depth map with high spatial resolution using very low photon counts is difficult using traditional estimation methods.

*3D imaging using active illumination:* Imagers based on the time-of-flight (TOF) principle modulate the intensity of light to vary as a function of time. Some examples are homodyne TOF sensing [6], pulsed TOF cameras [7], and picosecond laser radar systems [8]. While these require specialized detectors, their advantages are millimeter-resolution sensing and high photon efficiency due to the optical illumination source having low duty cycle. Sensors that spatially modulate the light include speckle decorrelation imaging, structured light, and active stereo imaging [9]. These can be implemented using standard CMOS detector arrays, their disadvantages are centimeter-resolution sensing and low photo efficiency due to the optical illumination source having high (often 100%) duty cycle. Among the aforementioned systems, TOF imagers using SPAD detectors have the highest photon efficiency. Our proposed framework further improves the photon efficiency of SPAD imagers, which translates to lower optical power and lower system bandwidth without sacrificing imaging quality.

This material is based upon work supported by the National Science Foundation under Grant No. 1161413, a Qualcomm Innovation Fellowship, a Microsoft PhD Fellowship, and a Samsung Scholarship.

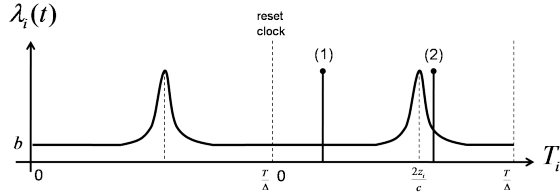


Fig. 1. Rate function of inhomogeneous Poisson process combining desired scene response and background noise signal. Individual single photon detections from noise (1) and signal (2) are shown.

## II. IMAGING SETUP

The light source and SPAD are co-located. Scene patches are indexed by  $i \in \{1, 2, \dots, M\}$ . The reflectance of patch  $i$ , denoted  $a_i$ , includes the effect of radial fall-off, viewing angle, and material properties. The distance to patch  $i$  is denoted  $z_i$ .

We use an intensity-modulated light source with pulse shape  $s(t)$  and pulse repetition interval  $T$ . We assume that, to avoid distance aliasing,  $T > 2z_{\max}/c$ , where  $z_{\max}$  is largest scene depth and  $c$  is the speed of light. We also assume  $T_w \ll 2z_{\max}/c$ , where  $T_w$  is the root mean square (RMS) pulse width, so that fine localization of depth is possible; for conventional ML estimation (MLE),  $T_w$  governs the achievable depth resolution [3].

In practice, every detected photon is time stamped within an accuracy of  $\Delta$ , called the *time bin size*. We assume  $\Delta \ll T_w$  so that the time bin size is not a significant impediment to fine depth resolution. In summary, the various time parameters satisfy  $\Delta \ll T_w \ll 2z_{\max}/c < T$ . When the detector records a photon arrival—referred to as a *click*—it becomes inactive for a small duration called the reset time or dead time.

For each raster position  $i$ , we repeatedly illuminate scene patch  $i$  until the first click. The discrete time bin  $T_i$  of the click (relative to the time that the most recent pulse was emitted) is measured to form the data set  $\{T_i\}_{i=1}^M$ .

## III. PROBABILISTIC MODEL

Illuminating scene patch  $i$  with intensity-modulated light pulse  $s(t)$  results in backscattered light intensity  $r_i(t)$ , which is modeled as  $r_i(t) = a_i s(t - 2z_i/c) + b_\lambda$ , where  $b_\lambda$  denotes the background light intensity at the operating wavelength. We denote by  $\eta$  the detector quantum efficiency, which is the fraction of photons that trigger electrical impulses. Accounting for the quantum nature of light, the function  $\eta r_i(t)$  is thus the rate of an inhomogeneous Poisson process observed at the SPAD, and photon arrivals are recorded per these statistics. The SPAD detector records these photon arrivals in time bins of size  $\Delta$ . It is also necessary to add the detector dark counts, which are modeled as an independent Poisson process with rate  $d$ . The observed inhomogeneous Poisson process has rate

$$\lambda_i(t) = \eta r_i(t) + d = \eta a_i s(t - 2z_i/c) + (\eta b_\lambda + d).$$

By defining  $b = b_\lambda + d/\eta$ , we see that  $\eta$  is an overall multiplicative factor that can be set to 1 without loss of

generality. Repetitions until the first click create a periodic rate function as shown in Figure 1.

*Derivation of distributions:* The depth of patch  $i$  is probabilistically encoded in the bin index  $T_i$ , which is a discrete random variable. Let  $N_{t_1, t_2}$  be the number of photon detections in the interval  $[t_1, t_2]$ . For the moment, assume  $b = 0$ . Then, the probability mass function (PMF) for  $T_i$  satisfies

$$\begin{aligned} p_{T_i}(k_i; z_i, a_i) &\propto \Pr[N_{0, (k_i-1)\Delta} = 0] \Pr[N_{(k_i-1)\Delta, k_i\Delta} \geq 1] \\ &= \exp\left\{-a_i \int_0^{(k_i-1)\Delta} s\left(x - \frac{2z_i}{c}\right) dx\right\} \\ &\quad - \exp\left\{-a_i \int_0^{k_i\Delta} s\left(x - \frac{2z_i}{c}\right) dx\right\}, \end{aligned}$$

where the dependence of the PMF on the parameters  $z_i$  and  $a_i$  has been emphasized. When  $\Delta$  is infinitesimally small, we observe a continuous-valued arrival time with probability density function (PDF) satisfying

$$f_{T_i}(v_i; z_i, a_i) \propto s\left(v_i - \frac{2z_i}{c}\right) \exp\left\{-a_i \int_0^{v_i} s\left(x - \frac{2z_i}{c}\right) dx\right\}.$$

We would like to know how  $f_{T_i}(v_i; z_i, a_i)$  behaves under low light levels (small  $a_i$ ). We observe the following.

**Theorem 1.** *Let  $U$  be a continuous random variable with PDF  $f_U(u; z_i) = s(u - 2z_i/c) / \int_0^T s(x - 2z_i/c) dx$ , i.e., the distribution obtained by shifting and normalizing the emitted pulse waveform. Then  $\lim_{a_i \rightarrow 0^+} f_{T_i}(u; z_i, a_i) = f_U(u; z_i)$ .*

Because of Theorem 1, we can use the approximation

$$f_{T_i}(v_i; z_i, a_i) \stackrel{\text{small } a_i}{\approx} \frac{s(v_i - 2z_i/c)}{\int_0^T s(x - 2z_i/c) dx}. \quad (1)$$

Under this approximation, our estimate of  $z_i$  is valid when the light intensity at the detector is low, as would be the case when photon efficiency is important.

*Signal vs. noise:* In order to mitigate the background noise, one may try to estimate background level  $b$  through calibration and directly model the likelihood as

$$f_{T_i}(v_i; z_i, a_i) \stackrel{\text{small } a_i, b}{\approx} \frac{a_i s(v_i - 2z_i/c) + b \mathbb{1}_{[0, T]}(v_i)}{\int_0^T (a_i s(x - 2z_i/c) + b) dx},$$

where  $\mathbb{1}_{[0, T]}(\cdot)$  is the indicator function for the interval  $[0, T]$ . However, including the uniform distribution in the likelihood leads to a nonconvex MLE problem. Thus, we instead rely on a censoring scheme that identifies and rejects background photons. We use  $\mathcal{S}$  to denote the set of patch indices that have signal photons, and PPPI includes a step to estimate  $\mathcal{S}$ .

## IV. NOVEL IMAGE FORMATION

Our proposed computational 3D image formation method uses the regularized ML estimation technique. We combine physically-realistic probabilistic signal models derived in the previous section with regularization using sparsity. With this approach, we achieve significant improvements over pointwise ML and over traditional denoising as post-processing applied

without a detailed physical model. The image formation happens in 2 steps.

**1. Noise detection identification:** Noise events are generated from a homogeneous Poisson process with constant rate, resulting in detection times with the uniform distribution on  $[0, T]$ . In contrast, signal photons are generated by a time-inhomogeneous Poisson process whose rate is proportional to the backscattered waveform (concentrated in a time duration  $T_w \ll T$ ). Thus, combining the two sources of detected photons, we have approximately a mixture distribution combining a uniform distribution and an impulsive distribution.

The rank-ordered absolute differences (ROAD) statistic was introduced in [10] to provide a hypothesis test to separate an impulsive distribution from a higher-variance distribution. For patch  $i$ , the ROAD statistic uses the neighborhood  $N(i)$  of size 8 containing the immediate neighbors. It is defined as

$$\text{ROAD}(i) = \underset{i_1, i_2, i_3, i_4 \in N(i)}{\operatorname{argmin}} \left( \sum_{k=1}^4 |v_i - v_{i_k}| \right),$$

where  $i_1, i_2, i_3, i_4$  must be distinct. Intuitively,  $\text{ROAD}(i)$  is a measure of how much the value  $v_i$  deviates from its neighbors, with only the 4 nearest values used so that the effect of the high-variance component in the neighboring values is minimized. A high value for  $\text{ROAD}(i)$  indicates that few neighboring patches have nearby values, so the detection at patch  $i$  is likely to be due to a noise detection. To obtain our estimated signal patch set  $\mathcal{S}$ , we thus reject detections likely to be due to noise using the ROAD statistic as follows:

$$\begin{cases} i \in \mathcal{S}, & \text{if } \text{ROAD}(i) < 4(2T_w); \\ i \notin \mathcal{S}, & \text{if } \text{ROAD}(i) \geq 4(2T_w). \end{cases}$$

More generally, if we estimate using  $L$  independent single-photon detections per patch, then the binary hypothesis testing will be based on the statistic

$$R(i) = \frac{1}{L} \sum_{k=1}^L \text{ROAD}^{(k)}(i),$$

where  $\text{ROAD}^{(k)}(i)$  is the ROAD statistic at patch  $i$  computed from the  $k$ th detection at each patch.

**2. Depth estimation:** Once noisy pixels are identified, we can readily formulate the negative log-likelihood function for estimation of depth at patch  $i \in \mathcal{S}$  from censored continuous photon arrival times  $\{v_i\}_{i \in \mathcal{S}}$  as

$$\mathcal{L}_z(z_i; v_i) = -\log f_U(v_i; z_i) \equiv -\log s(v_i - 2z_i/c),$$

where the final equivalence omits an additive constant that does not depend on  $z_i$ . Our framework allows the use of arbitrary pulse shapes, but many practical pulse shapes, like Gaussian and lopsided pulse shapes, are well approximated as  $s(t) = e^{-\bar{f}(t)}$ , where  $\bar{f}(t)$  is a convex function in  $t$ . In such cases,  $\mathcal{L}_z(z_i; v_i) = \bar{f}(v_i - 2z_i/c)$  is a convex function in  $z_i$ . Our regularized ML depth estimate is obtained from filtered data  $\{v_i\}_{i \in \mathcal{S}}$  by solving the following constrained optimization problem:

$$\underset{\substack{Z = \{z_1, \dots, z_M\} \\ z_{\min} \leq z_i \leq z_{\max}}}{\operatorname{argmin}} \sum_{i \in \mathcal{S}} \mathcal{L}_z(z_i; v_i) + \beta \|Z\|_{\text{TV}}, \quad (2)$$

where  $\|\cdot\|_{\text{TV}}$  is the total variation seminorm. Note that (2) is essentially an image inpainting problem.

## V. NUMERICAL EXPERIMENTS

We simulated low-light SPAD-based imaging for the scene shown in Figure 2(a) to demonstrate the depth accuracy of our PPPI framework. For every patch, we model the scene amplitude as  $a_i = A/z_i^2$ , where  $A$  is a constant. From patch to patch, the probabilities of observing background photons will thus not necessarily be the same. Our transmitted pulse is modeled with a generalized Gaussian function defined as  $s(t) \propto \exp(-(|t|/a)^p)$ , where  $p$  and  $a$  control the concavity and the spread of the pulse, respectively. The generalized Gaussian family includes Gaussian ( $p = 2$ ) and near-uniform (large  $p$ ) functions. Since generalized Gaussian functions are log-concave, our regularized depth estimator is obtained by solving a simple convex program. The regularization parameter  $\beta$  in (2) was chosen to maximize PSNR.

Figure 2 shows depth map recovery results only using 1 photon detection per pixel. The setup parameters were  $p = 3$ ,  $a = 10$ , and  $T = 200a$ . The background light level was set such that the probability of a detected photon being a background photon, on average over the whole field of view, was 0.32. Our computational imaging method has high PSNR gain over well-known denoising algorithms, which are post-processing methods on the MLE result. Median filter denoising successfully removes salt-and-pepper noise but it fails in preserving the fine edge features of the scene. The state-of-the-art BM3D denoising [11] seems to reject most of the noisy data but we see that it introduces a bias in the background and thus leading to a moderate PSNR gain. Figure 3 zooms in to several regions of recovered depth maps. We clearly see that our computational imager can finely resolve object gradients and even make out objects in the far field region, where background photons are more likely to be detected.

Figure 4 shows the high photon efficiency of PPPI compared to the traditional pixelwise MLE method. In this setup, we image the same scene with chess pieces but the parameters used are  $p = 2$ ,  $a = 10$ , and  $T = 200a$ . Also, the background light level was set such that the mean probability of a detected photon being a background photon is approximately 0.2. The black dashed lines show that at least 80 ppp is required for MLE to achieve the same PSNR that our computational imager achieves using 1 detected ppp.

## VI. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a new 3D imaging framework that enables robust scene depth acquisition with very few photon detections under high background light illumination. Our method accurately models the photon arrival statistics and exploits the piecewise smoothness of depth maps to form accurate images. We demonstrate the high photon efficiency of our computational imaging method over the traditional ML estimator. We also show that our method is more successful in preserving high frequency details and small features compared to popular image denoising algorithms.

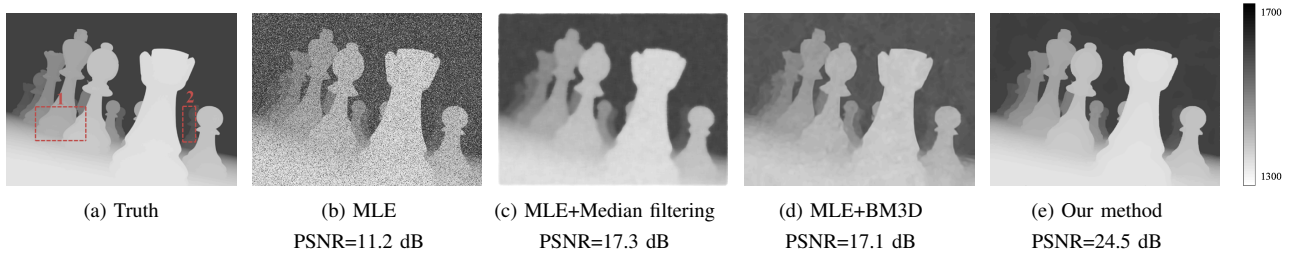


Fig. 2. Depth map results using only one detected photon per pixel.

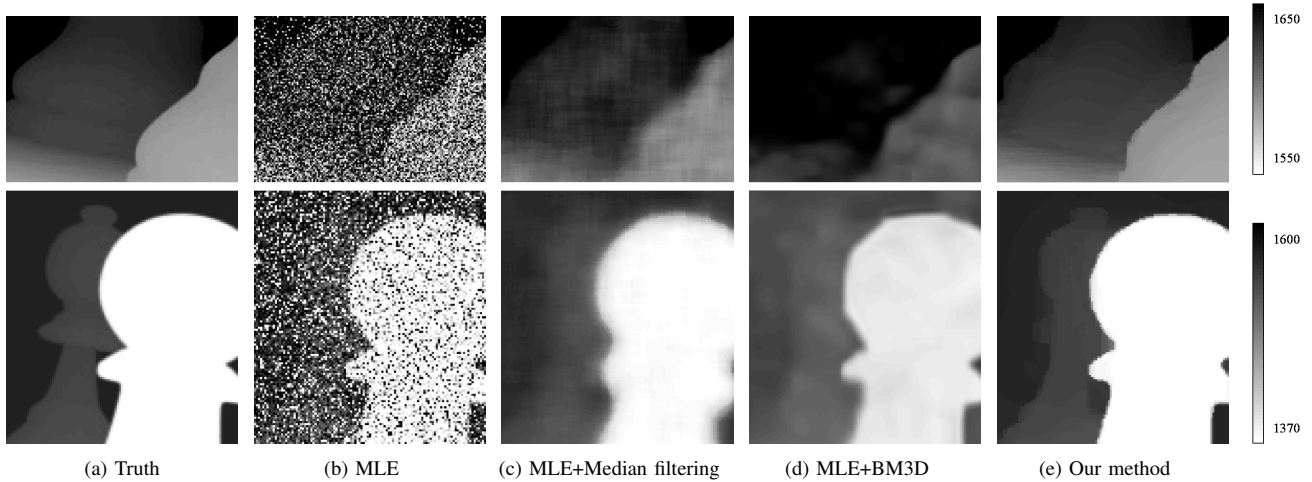


Fig. 3. Depth map of chess pieces (zoomed in boxes from Figure 2(a)). Our method clearly identifies gradients and edges of the chess pieces.

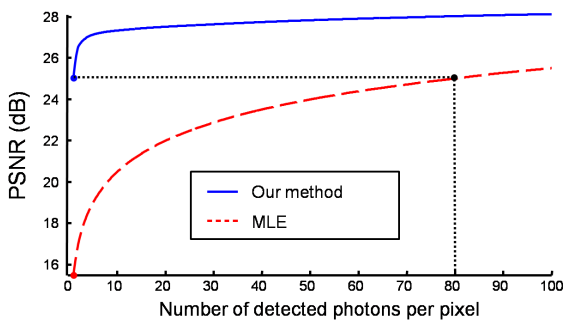


Fig. 4. PSNR vs. ppp plot showing high photon efficiency of PPPI (blue, solid) over pixelwise ML estimation (red, dashed) with contribution from background light.

Although this paper mainly investigates imaging at low light levels, our approach may also extend to optical imaging using photon-counting detectors at high optical flux. Accurate likelihood modeling under arbitrary flux levels to estimate depth is thus of future interest. Furthermore, the PPPI framework extends to intensity imaging as well.

#### ACKNOWLEDGMENTS

The authors thank Dheera Venkatraman, Franco N. C. Wong, and Jeffrey H. Shapiro for their contributions to the modeling and experiments.

#### REFERENCES

- [1] A. Kirmani, D. Venkatraman, A. Colaço, F. N. C. Wong, and V. K. Goyal, "High photon efficiency computational range imaging using spatio-temporal statistical regularization," in *Proc. CLEO*, San Jose, CA, Jun. 2013, paper QF1B.2.
- [2] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "Sparsity-regularized photon-limited imaging," in *IEEE Int. Symp. Biomed. Im.: From Nano to Macro*, Rotterdam, Apr. 2010, pp. 772–775.
- [3] R. M. Gagliardi and S. Karp, *Optical Communications*. Wiley, 1976.
- [4] G. Buller and A. Wallace, "Ranging and three-dimensional imaging using time-correlated single-photon counting and point-by-point acquisition," *IEEE J. Sel. Top. Quantum Electron.*, vol. 13, no. 4, pp. 1006–1015, Jul.–Aug. 2007.
- [5] D. L. Snyder, *Random Point Processes*. New York, NY: John Wiley & Sons, 1975.
- [6] S. B. Gokturk, H. Yalcin, and C. Bamji, "A time-of-flight depth sensor — system description, issues and solutions," in *Proc. Conf. Comput. Vis. Pattern Recog. Workshop*, 2004, p. 35.
- [7] G. J. Iddan and G. Yahav, "3d imaging in the studio (and elsewhere)," in *Proc. SPIE*, vol. 4298, 2001, pp. 48–55.
- [8] A. McCarthy, R. J. Collins, N. J. Krichel, V. Fernández, A. M. Wallace, and G. S. Buller, "Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting," *Appl. Optics*, vol. 48, no. 32, pp. 6241–6251, Nov. 2009.
- [9] S. B. Kang, J. A. Webb, L. C. Zitnick, and T. Kanade, "A multibaseline stereo system with active illumination and real-time image acquisition," in *Proc. 5th Int. Conf. Comput. Vis.*, 1995, pp. 88–93.
- [10] R. Garnett, T. Huegerich, C. Chui, and W. He, "A universal noise removal algorithm with an impulse detector," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1747–1754, Nov. 2005.
- [11] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, 2007.