This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

Ghosts in the Point Clouds: De-glaring LiDAR in the Transient Domain

Avery Gump* Connor Henley* Sungjin Cheong Akarsh Prabhakara Mohit Gupta
University of Wisconsin–Madison, Madison, WI, USA
{gump, connor.henley, sungjin.cheong, aprabhakara, mgupta37}@wisc.edu

wisionlab.com/project/deglaring-LiDAR

Abstract

Modern LiDARs are rapidly transitioning from bulky, mechanically scanned systems to ultra-compact, low-cost, solid-state arrays. This miniaturization—while enabling scalability, affordability, and camera-like data structures—introduces a new and severe failure mode: internal-multipath glare. When light from a bright or retroreflective surface reflects and scatters within the LiDAR, light that should reach a single pixel spreads across the pixel array. The resulting artifacts create phantom objects, obscure real ones, and produce safety-critical "ghosts in the point clouds." This paper introduces a physically grounded sensing model and algorithmic techniques for addressing this effect. We show that internal glare can be represented as a linear, scene-independent operator—the Transient Glare Spread Function (TGSF)—acting on the transient measurements. Building on this model, we develop a training-free approach that operates on low-level LiDAR detections (or echoes) prior to point-cloud formation, leveraging knowledge of the glare spread function to reason about the likelihood of each detection arising from glare. The resulting approach is compatible with existing LiDAR signal-processing pipelines, and deployable on unmodified commercial sensors. Using experiments with real single-photon LiDAR hardware, we demonstrate substantial suppression of severe glare artifacts while preserving true scene structure.

1. A Changing LiDAR Landscape

LiDAR technology is undergoing a rapid and fundamental transformation. The field is moving away from legacy scanning-based systems—bulky, expensive, and mechanically complex—toward compact, solid-state architectures that are low-cost and scalable. This transition, driven in large part by the emergence of high-resolution *Single-Photon Avalanche Diode (SPAD)* arrays [4], promises to transform LiDAR from a "necessary evil" into a reliable and inexpensive modality, much like "just another, but more informative, camera." Furthermore, dense pixel arrays in future LiDARs will make their data resemble camera imagery in structure and scale, potentially merging camera and LiDAR processing pipelines.

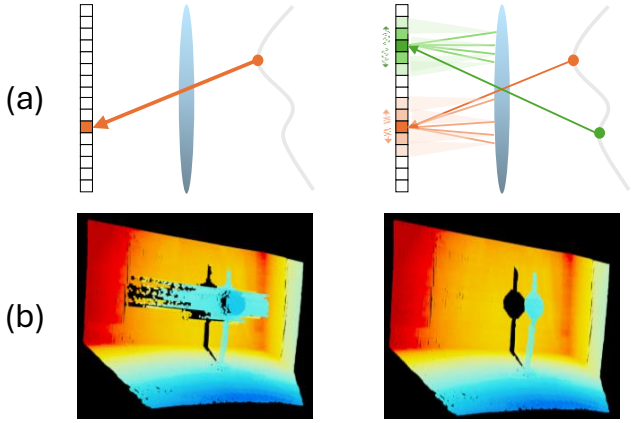


Figure 1. (a) For an ideal LiDAR sensor, there is a one-to-one correspondence between scene point and sensor pixel. (b) "Camera-fication" of LiDAR leads to glare from internal optical scattering, multiple internal reflections, and electrical crosstalk.

This *camera-fication* of modern LiDARs comes with a cost. As in conventional cameras, modern LiDAR arrays view the world through shared optical elements. When a bright return enters the system, light that should ideally be focused onto a single pixel can undergo multiple internal reflections between the lens and sensor, and can scatter within the optical assembly itself, redistributing itself across a wider neighborhood of pixels. The result is an increasingly critical phenomenon we refer to as **internal-multipath glare** (also called *blooming* in the LiDAR community).[1] The redistribution creates false or "phantom" structures while suppressing true returns from real surfaces. This results in LiDARs *seeing what is not there* (e.g., a phantom wall materializing on a freeway due to overhead signs that cause an autonomous vehicle to *brake*), and simultaneously *missing what is there* (e.g., a child standing beneath a retroreflective stop sign could be obscured). Such

*Equal contribution

†This research was supported by NSF CAREER Award 1943149, NSF grant CNS-2107060, NRT Award Number 2152163, and ONR grant N000142412155. We thank Sacha Jungerman for helpful discussions.

[1] Unlike glare in passive imaging [25, 28]—typically driven by strong external sources such as sun or headlights in autonomous driving or bright stars in astronomy—LiDAR glare is self-induced by intense returns.

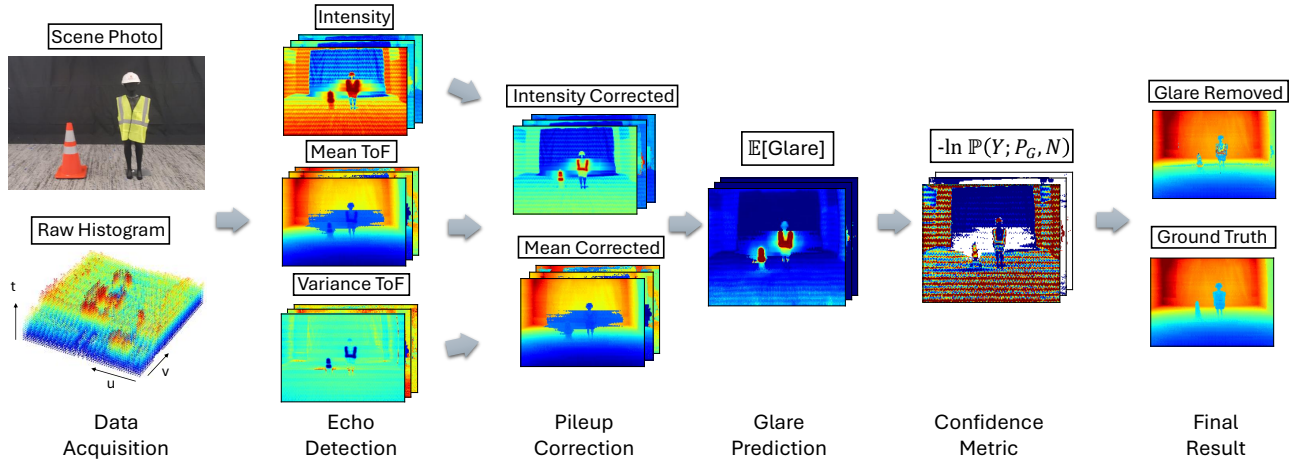


Figure 2. Overview of our glare mitigation method. First, raw histograms are processed to extract intensity, mean time-of-flight (ToF), and ToF variance of several echoes per pixel. Second, a pileup correction method corrects the intensity and ToF of the brightest returns, which induce glare. Third, a glare model predicts the glare contribution to each echo. Finally, a confidence metric quantifies the likelihood that each echo contains signal in excess of the predicted glare. We choose the highest confidence echo at each pixel to form a depth map.

'ghosts in the point clouds' are safety-critical failure modes. One of the most common culprits are retroreflectors (road signs, license plates, lane markers, safety vests, and construction cones). These objects are deliberate and ubiquitous components of modern infrastructure; constructing a realistic scenario *without* glare sources may indeed be harder than constructing one with them.

Why has this failure mode remained largely invisible to the research community? Historically, legacy LiDARs avoided severe optical crosstalk by design. Commercial scanning systems such as the Velodyne HDL-64e grouped detectors behind separate optical paths and—critically—fired only one detector per group at any given time.[2] This combination of dedicated optics and scanning mechanisms effectively suppressed internal reflections, but at significant cost: large form factors, spinning mechanisms with higher cost/complexity, sparse point clouds, motion-induced artifacts, and limited resolution. Indeed, the severity of blooming was such that prior LiDAR generations adopted fundamentally sub-optimal hardware architectures to avoid it. Consequently, widely used datasets and benchmarks [1, 3, 7], captured using such systems, exhibit minimal glare. The landscape is now changing. Modern solid-state LiDARs unlock high sensitivity, fine temporal resolution, and high-resolution arrays at low costs. However, these architectures alter internal light transport and acquisition, making internal glare not a corner case but a frequent and severe artifact.

Internal glare breaks a foundational assumption underlying LiDAR processing: the one-to-one correspondence between a pixel and a single scene point (Fig. 1). Once light from a bright return spreads across pixels, downstream steps (peak-finding, denoising, and point-cloud processing) operate on corrupted data. Waiting to address glare until after point cloud formation is often too late: the artifacts, becoming entangled with geometry, are difficult to separate from true scene elements, and retroreflector-driven ghosts can appear more confident than the real objects.

We argue that internal glare should be addressed early in the sensing pipeline, *before* point-cloud formation. Modern single-photon (SP) LiDARs [4, 14, 15] capture rich temporal information in the form of transient histograms. In this measurement space, the physics of internal glare can be modeled as a *linear, scene-independent operator* acting on an ideal, glare-free transient histogram. We refer to this operator as the *Transient Glare Spread Function (TGSF)*, which describes how photons from a single pixel are redistributed across other pixels and neighboring time bins through internal reflections and scattering.[3]

Building on this observation, we develop LiDAR deglare algorithms that operate directly on low-level temporal peak detections (or *echoes*) derived from transient histograms, prior to point-cloud formation (Fig. 2). We leverage knowledge of the GSF as a forward model to reason about the likelihood that each detected echo arises from internal glare. This allows us to identify and suppress glare-induced detections while preserving true scene returns. Importantly, the method integrates naturally with standard LiDAR signal-processing pipelines and can be deployed on unmodified, commercially available sensors.

A key challenge in this regime is that strong

[2]For example, in the Velodyne HDL-64e, two groups of 32 detectors each share separate optics, and only one detector in each group is active at a time [29]. This design choice was motivated explicitly by the need to prevent blooming when encountering retroreflectors.

[3]In practice, the temporal extent of this spreading is typically small, and the effect can be treated as predominantly spatial; thus, we refer to the operator as the *Glare Spread Function (GSF)* in the rest of the paper.

glare—particularly from retroreflectors—is typically accompanied by severe nonlinear distortion due to photon pileup [6, 9, 9, 11]. The proposed approach therefore jointly addresses both internal glare and pileup effects, and is guided by the following principles:

- **Generality.** A single method applies across retroreflector types, orientations, sizes, counts, and scene configurations—without requiring explicit detection or modeling of glare sources.
- **Training-free.** The method does not require large-scale datasets, making it well-suited to single-photon LiDAR settings where labeled data is scarce.
- **Compatibility with downstream processing.** The method integrates seamlessly into existing LiDAR signal-processing pipelines prior to point-cloud formation, and can propagate confidence measures alongside each echo.
- **Efficiency.** The approach is computationally lightweight.

We demonstrate, using a commercial SP-LiDAR evaluation kit, that internal glare produces severe artifacts that standard pipelines do not remove—phantom returns and "ghost" objects that occlude real ones. Across scenes containing diverse and multiple retroreflectors, our method substantially suppresses these effects while preserving true structure, even in the presence of strong pileup distortion.

Contributions. This paper makes four key contributions:

- Identification of *internal-multipath glare* as a critical failure mode for next-generation solid-state LiDARs.
- A transient-domain formulation of glare that accounts for both internal multipath and nonlinear pileup distortion arising from high-intensity returns.
- A training-free algorithm that leverages knowledge of the glare spread function to estimate the likelihood that each LiDAR detection is induced by glare.
- A practical approach that operates prior to point-cloud formation, and integrates with standard LiDAR pipelines.

2. Related Work

Glare in passive imaging and astronomy. In conventional cameras, glare (often termed *veiling glare*) has been modeled and mitigated using linear optical spread models and inverse formulations. Talvala *et al*. [25] measured a glare spread function and applied deconvolution for glare removal. Glare has also been studied in astronomical imaging, which faces extreme dynamic range [28]. Optical filtering has been explored to improve the conditioning of the glare spread function [21]. As in these works, we use a linear GSF to model glare formation, but also consider the transient nature of glare in LiDAR, as well as nonlinear pileup effects unique to SP-LiDAR.

Glare in LiDAR. Work specifically addressing LiDAR internal glare is sparse. Recent learning-based approaches operate on point clouds to detect or correct blooming artifacts [5, 27], typically relying on synthetic training data and strong assumptions about reflector geometry. Because these methods act *after* point-cloud formation, artifacts are already entangled with geometry, and such approaches often struggle to generalize across reflector types, layouts, and intensities. By contrast, we use a glare model that depends only on internal sensor characteristics. As such, our method is training-free and generalizes to arbitrary scenes.

Electrical crosstalk in SPAD arrays, where secondary photons generated during an avalanche re-trigger neighboring pixels [2, 20, 26], has also been studied. However, such electrical crosstalk is typically confined to a small neighborhood of adjacent pixels, and modern devices largely mitigate this effect using deep trench isolation. Internal glare on the other hand arises from optical reflections and scattering between the SPAD array and shared optics, and can spread over large portions of the sensor, producing spatially extended artifacts.

Computational single-photon imaging. A rapidly growing body of work addresses single-photon LiDAR artifacts orthogonal to glare: pileup modeling and mitigation [9, 11], probabilistic inference and denoising for photon-limited histograms [8, 14–16, 18, 24], and compression/efficient streaming of transient data [10, 12, 22, 23]. These advances are complementary to our goals. Our method for identifying and suppressing glare-induced returns is compatible with existing pileup correction, denoising, peak-finding, compression, and downstream learning—suggesting a modular *photon-processing pipeline* for next-generation LiDAR.

3. Method

In Sec. 3.1 we propose a transient-domain glare model for glare in "camera-fied" LiDAR sensors. In 3.2 we describe our method for correcting photon pileup, an artifact that plagues extremely high-flux signals measured by SP-lidars, and usually co-occurs with glare. In 3.3, we review our method for mitigating glare in two steps: first, by predicting glare contributions to each detected echo; second, by determining the likelihood that each echo contains a true (non-glare) signal. Our full pipeline is visualized in Fig. 2.

3.1. Glare in the Transient Domain

Veiling glare in optical imaging systems arises from a variety of optical processes, including multiple reflections between lens and image sensor surfaces, and diffuse scattering from lens surfaces and aperture edges [13]. In photon-counting lidar systems, electroptical processes such as photon re-emission due to hot carriers relaxation can also produce a glare-like effect, often referred to as "crosstalk" [2, 20]. Several sources of glare are depicted in Fig. 1a,b. In spite of its myriad causes, glare formation is a linear process, in that the intensity of the glare signal responds linearly to the intensity of incident light.

We consider a camera-like, direct-time-of-flight (dToF) LiDAR sensor that consists of a time-resolving detector array positioned behind a lens in an imaging configuration,

and a transmitter that illuminates the scene with pulsed laser light in an arbitrary spatial pattern (e.g. flash, line scanned, grid of dots, etc.). The raw measurement captured by dToF cameras is a three-dimensional datacube that resolves the spatial (pixel) position of incident light (u, v) as well as its time of flight t. This means that, unlike in regular photography, glare observed by a dToF camera has a transient aspect—the time that glare is observed depends on the depth of the bright surface that produces it.

We define y as the transient intensity of light incident upon the image sensor that has been corrupted by glare and noise, and x as the "true" uncorrupted transient intensity. The process of glare formation is governed by the six-dimensional *transient glare spread function* (TGSF) $a(u, v, t, u', v', t')$, which quantifies the linear response of the measured datacube y at coordinate (u, v, t) to the uncorrupted intensity of x at (u', v', t'). Altogether the measurement process can be described by the following integral equation

$$y(u,v,t) = \iiint a(u,v,t,u',v',t')x(u',v',t')du'dv'dt' + \eta(u,v,t). \tag{1}$$

Here $\eta(u, v, t)$ represents additive noise attributable to ambient background light, detector dark counts, and detector afterpulsing [17]. We note that the TGSF only accounts for the effects of glare. It does not account for other effects such as spatial blur due to defocus, or temporal blur due to laser pulsewidth or detector timing jitter. Hence, if these other effects are present in the observed y, they will also be present in the uncorrupted intensity x.

In theory, scattering within the receiver optics that produces glare alters the path lengths of measured photons, resulting in a time delay that must be accounted for in the TGSF. In practice, these time delays are typically very small relative to the time resolution of the lidar sensor, and can be safely ignored. As a consequence, the TGSF reduces to a *time-independent* glare spread function (GSF), equivalent to the GSF used for glare removal in passive photography [25, 28]. The measurement model can then be updated:

$$y(u,v,t) = \iint a(u,v,u',v')x(u',v',t)du'dv' + \eta(u,v,t). \tag{2}$$

3.2. Correcting Pileup Distortion

Glare and Pileup The time-independent GSF model of Eq. 2 suggests a straightforward approach to glare removal: each time slice in the measured datacube can be treated as an independent still image and de-glared separately using photographic de-glare methods [25, 28]. While appealing in theory, this approach fails in practice due to the phenomenon of *photon pileup*, which is inherent to SPAD-based LiDAR and almost always co-occurs with glare in such systems.

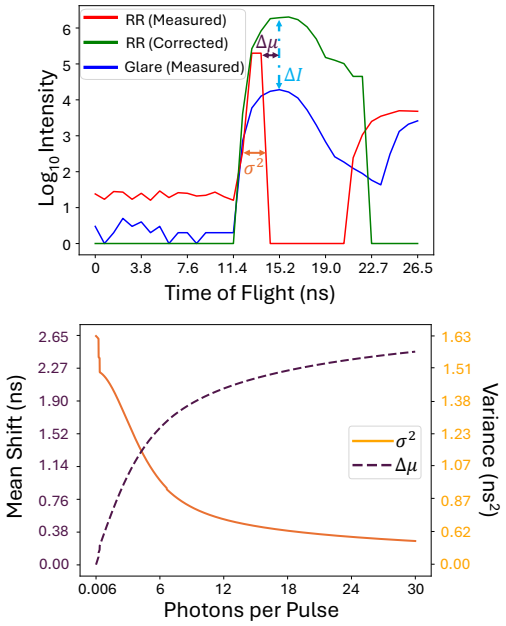


Figure 3. (Top) Pileup distortion shifts mean and intensity of histogram corresponding to intense retroreflector (RR) return. RR return induces fainter glare return at a second pixel, which is not affected by pileup. (Bottom) Mean shift and variance of pileup distorted echoes corresponding to a range of incident photon fluxes.

Photon pileup occurs because SPADs must switch off for several nanoseconds after each photon detection to reset—a period known as the *dead time* [17]. When SPADs are subjected to very high fluxes, for which the probability of multiple photon arrivals in a single dead-time period is high, early-arriving photons censor late-arriving photons, distorting the measured pulse shape relative to the true incident waveform. Glare interference is typically several orders of magnitude fainter than the direct returns that induce it. As such, problematic glare, in general, only co-occurs with extremely high-flux direct returns from specular, retroreflective, or very-close range surfaces.

Photon pileup has the practical effects (depicted at top of Fig. 3) of reducing measured flux relative to true incident flux, and shifting the measured signal earlier in time (producing *range-walk error*). Using Eq. 2 to naively predict glare from pileup corrupted histograms would yield predictions that were underestimated and temporally misaligned relative to the true glare signal.

Correcting for pileup at the level of raw histograms can be achieved using Coates's method [6], but Coates's method saturates at fluxes as low as 1-10 incident photons per pulse [11], and as such is inadequate for returns produced by extremely reflective targets like retro-reflectors. Higher dynamic range pileup correction can be achieved if one estimates the true time of flight and intensity of a finite num-

ber of returns in each histogram [11, 19, 23] rather than fully "undistorting" the raw histogram itself. For this reason, we argue that glare removal is most effectively addressed among low-level detections—that is, the finite number (e.g. 1, 2 or 3) of returns or *echoes* reported per-pixel, per-frame by many commercial LiDAR sensors.

Moments-based Pileup Correction We detect echoes by convolving each raw histogram with a matched filter that is a replica of the transmitted waveform, and then identifying the three highest peaks in the filtered histogram. We then position a fitting window around each peak and estimate the first three moments of photon arrival times (number of photons, mean time-of-flight, and time-of-flight variance) in each window. We additionally estimate the level of background noise counts by averaging counts in a noise window that reliably contains only background detections.

If the measured echo intensity is high enough (e.g. $> .05$ photons per pulse) to produce significant pileup distortion, we estimate the echo's true time of flight and intensity using a moments-based method similar to that proposed by Sheehan et al.[23]. As the true echo intensity increases and pileup becomes more severe, photons bunch up into fewer timing bins—reducing the *variance* of the time-of-flight distribution—and shift to earlier arrival times—shifting the *mean* of the distribution. We use a photon pileup model to determine the relationship between the true echo intensity, and the measured echo's variance and mean shift. Our model requires knowledge of waveform shape, detector dead-time, and the rate of background counts. Moments estimated with the model are shown in Fig. 3, and theoretical details are provided in the Supplement.

We use the pileup model to compile lookup tables that are queried at measurement time. First, the true echo intensity is obtained from the measured variance via an inverse lookup. This estimated intensity is then used to determine the mean shift caused by pileup. The estimated mean shift is subtracted from the measured mean to obtain the corrected, true mean time-of-flight, which is in turn used to compute the target range.

3.3. Identifying Glare-induced Echoes

Our glare suppression method operates in two steps. First, we use a glare model to predict the expected glare intensity at the time of flight and spatial position of each detected echo. Second, we compare the predicted glare intensities to the measured echo intensities to determine the likelihood that each echo was induced by glare.

Echo-space Glare Model We consider a dToF LiDAR camera with M pixels, and extract the intensity and time-of-flight of K echoes per pixel, per frame. The vector $\mathbf{y} \in \mathbb{R}_+^{M\times K}$ represents the pileup-corrected intensities (in units of photon counts) of measured echoes, and $\mathbf{t} \in \mathbb{R}_+^{M\times K}$ the pileup-corrected mean times of flight of those echoes. The vector $\mathbf{x} \in \mathbb{R}_+^{M\times K}$ represents the true intensity of each echo after the influence of glare is removed.

To reformulate the glare model of Eq. 2 in the discrete echo space, we integrate over the duration of the fitting window used for pileup correction

$$\begin{aligned} y_{uk} &= \int_{t_{uk}-\Delta/2}^{t_{uk}+\Delta/2} y(u,t)dt \\ &= \sum_{u'=1}^{M} a(u,u') \left[\int_{t_{uk}-\Delta/2}^{t_{uk}+\Delta/2} x(u',t)dt \right] + \eta_{uk}. \end{aligned} \tag{3}$$

Here y_{uk} is the measured intensity of echo k at pixel u, $y(u,t)$ and $x(u,t)$ are the glare-corrupted and true transient intensities, respectively, defined in Sec. 3.1, and $a(u,u')$ is the GSF. For compactness, the continuous, two-dimensional spatial coordinates (u,v) have been reduced to a discrete, one-dimensional coordinate u. The symbol Δ denotes the fitting window duration, and η_{uk} is the level of background noise in the fitting window. Here the window is centered on the echo mean, but it would be straightforward to incorporate a constant offset of the window center.

For a given echo at pixel u, the temporal fitting windows corresponding to echoes at *other* pixels may not overlap perfectly in time. We introduce a *temporal overlap function* $o(\delta t)$—defined as the correlation of the fitting window with the transmitted pulse shape—to quantify the fraction of the signal in another echo k' at pixel u', that contributes glare to echo k at pixel u, i.e.

$$o(\delta t) = \int_{-\infty}^{+\infty} \text{rect}\left(\frac{t}{\Delta}\right) \lambda(t-\delta t)dt \tag{4}$$

where $\delta t = t_{u'k'} - t_{uk}$ and $\lambda(t)$ is the shape of the transmitted waveform, normalized so that $\int_{-\infty}^{+\infty} \lambda(t)dt = 1$.

Eq. 3 can now be written in discrete form

$$y_{uk} = \sum_{u'=1}^{M} a(u,u') \left[\sum_{k'=1}^{K} o(t_{u'k'} - t_{uk}) x_{u'k'} \right] + \eta_{uk}. \tag{5}$$

The above model can be expressed as a linear system $\mathbf{y} = \mathbf{A}\mathbf{x} + \eta$, where entries of $\mathbf{A}$ correspond to the GSF modified by temporal overlaps. In theory, it is possible to invert this system to obtain the glare-free intensities $\mathbf{x}$ from the glare-corrupted intensities $\mathbf{y}$. However the matrix $\mathbf{A}$ depends on measurements of $\mathbf{t}$, which vary frame-to-frame. Inverting a unique linear system for each LiDAR frame would be computationally demanding. Instead, we approximate the glare intensity $\bar{g}_{uk}$ at pixel u, echo k with the operation:

$$\bar{g}_{uk} = \sum_{u'=1}^{M} a(u,u') \left[\sum_{k'=1}^{K} o(t_{u'k'} - t_{uk}) y_{u'k'} \right], \tag{6}$$

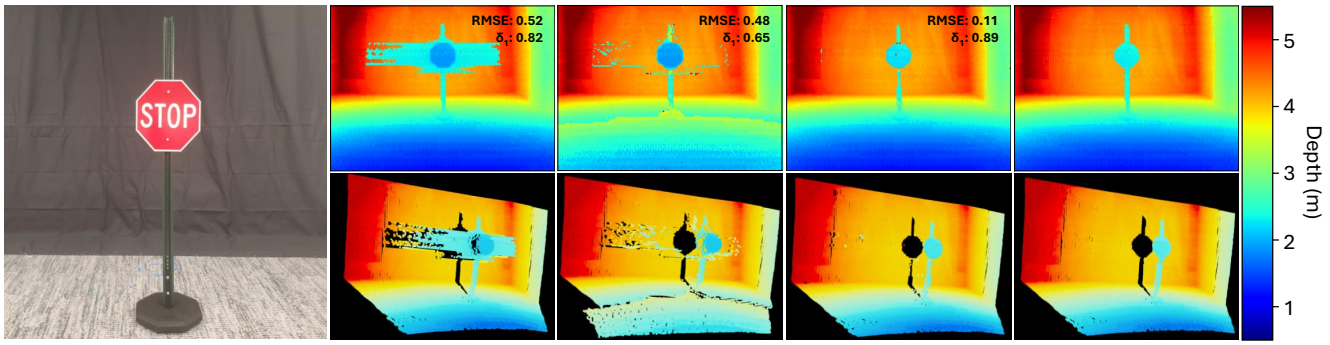


Figure 4. (a) Photograph of captured scene. (b) Measured, glare-corrupted depth. (c) Corrected depth using baseline, deep learning method [5]. (d) Corrected depth using our method. (e) Ground truth depth captured by covering retroreflectors with black construction paper. (insets). Root mean square error (RMSE ↓) in meters and δ_1 ↑ (see Eq. (9)) of depth maps relative to ground truth.

which is equivalent to the backprojection $\bar{\mathbf{g}} = \tilde{\mathbf{A}}^T\mathbf{y}$, where $\tilde{\mathbf{A}}$ is equal to $\mathbf{A}$ with diagonal entries set to zero.

Binomial Likelihood Confidence Metric We seek a confidence metric to assess the likelihood that a detected echo contains signal in excess of the predicted glare. We define Y as the number of photon detections associated with an echo *before* pileup correction. We assume Y follows a binomial distribution [4], with the number of trials N corresponding to the number of laser pulses. We define $P_G = G/N$ as the per trial probability of detecting a photon *if* the incident flux equals the predicted glare flux computed with Eq. 6. The parameter G is the expected number of *glare* photons expected in N trials[5]. The probability of detecting Y photons in N trials given a per-trial success probability P_G is

$$\mathbb{P}(Y; P_G, N) = \frac{N!}{Y!(N-Y)!} P_G^Y (1-P_G)^{N-Y}. \quad (7)$$

If Y is much greater than NP_G, then $\mathbb{P}(Y; P_G, N)$ will be small. Following this logic, we define the following confidence metric:

$$C = \begin{cases} -\ln \mathbb{P}(Y; P_G, N), & \text{if } Y \geq NP_G \\ 0, & \text{if } Y < NP_G \end{cases} \quad (8)$$

The confidence metric C acts as a threshold when detected counts lie below the expected counts from glare, and increases monotonically for counts in excess of the predicted glare.

The second-to-last column of Fig. 2 shows a map of confidence values computed for the brightest echo at each pixel, when observing a scene containing multiple retro-reflectors. We see that, in general, $C = 0$ (white pixels) when the brightest echo corresponds to glare, and $C > 0$ for non-glare echoes. In practice, when multiple echoes are detected per pixel, the confidence metric can be used to determine which echoes correspond to real surface returns, and which correspond to glare.

[4]This assumption is valid if the fitting window duration is shorter than the SPAD deadtime or, alternatively, if the probability of detecting more than one photon, per laser pulse, in the fitting window is small.

[5]For moderate to low fluxes, $G \approx g_{uk}$. For high glare flux where pileup is significant, G must be obtained from an inverse lookup table generated from a pileup model. See Supplement for details

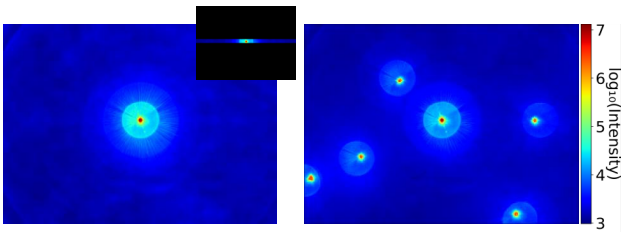


Figure 5. (a) Glare spread function (GSF) captured using infrared flashlight as light source. Here, source is at FOV center. (inset) Cropped, six-row GSF that accounts for synchronous row-scanning of LiDAR sensor. (b) Superposition of GSFs captured at multiple source positions demonstrates spatial variation of GSF.

4. Experiments

Experiments were conducted with a commercial SP-LiDAR evaluation kit (ADS6311, Adaps Photonics). ADS6311 is a dToF LiDAR array with 192x256 SPAD macro-pixels, that illuminates the scene with a vertically scanned, horizontally oriented line beam. Evaluations were run on a RTX 4090 (24 GB) with Ubuntu 22.04. We demonstrated our method on multiple scenes with varying scene complexity.

GSF Acquisition To measure the GSF, we turn off the LiDAR transmitter and use an IR flashlight at the same wavelength (940nm). This allowed us to isolate light to a small spot on the receiver array, and alleviated dynamic range constraints imposed by pileup distortion by spreading incident light across all timing bins. We capture the GSF for 49 light source positions in the LiDAR FOV, and use a normalized distance-weighted sum to interpolate to light source positions that were not measured. An example GSF with the light source at the FOV center is shown in Fig. 5.

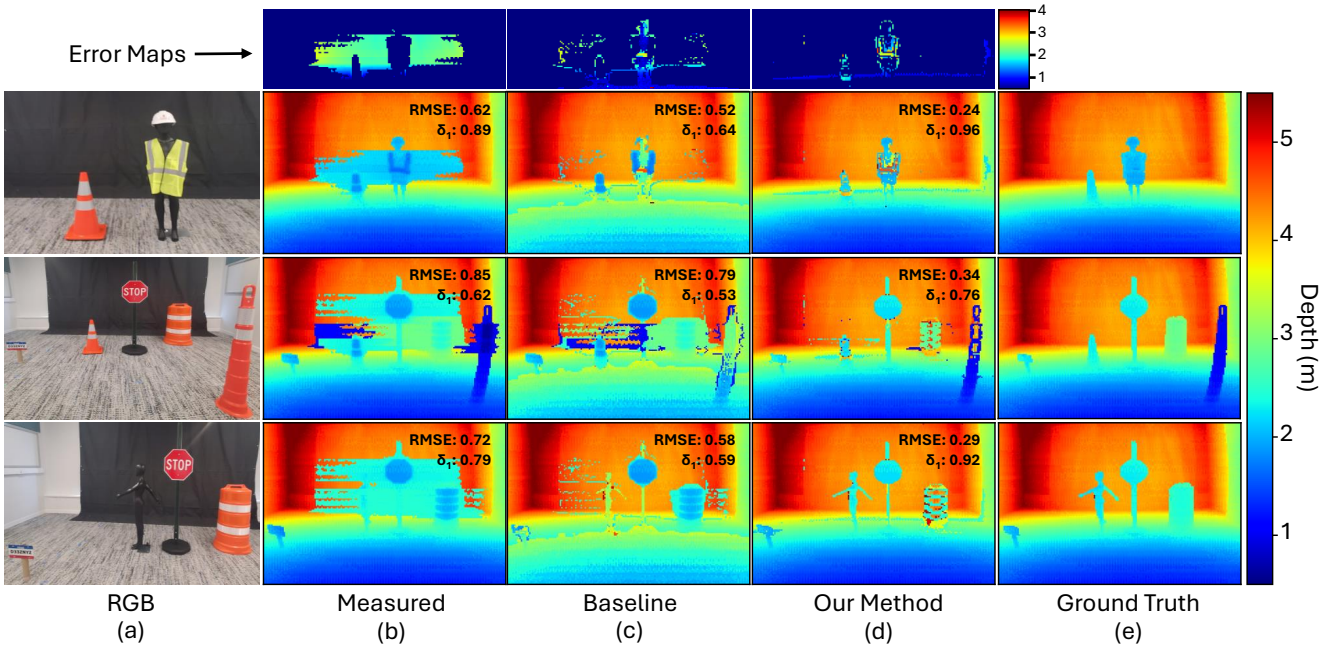


Figure 6. (a) Photographs of captured scenes (b) Measured, glare-corrupted depth maps (c) Depth maps corrected using baseline, deep learning method [5] (d) Depth maps corrected using our method (e) Ground truth depth maps captured by covering retroreflectors with black construction paper. (insets) Root mean square error (RMSE ↓) in meters and δ_1 ↑ (see Eq. (9)) of depth maps relative to ground truth

ADS6311 scans the scene vertically, switching on groups of six pixel rows in sync with the scan position of the transmitted beam. This means that during LiDAR measurements, glare is observed in six-row bands around the aggressor pixel. The banding is not observed in flashlight measurements because the flashlight is not scanned synchronously with sensor row activations. We re-create the banding effect by masking pixels outside a band that contains the light source (inset, Fig. 5).

Evaluation Setup To evaluate our method we used a number of different retroreflectors common on the road: a stop sign, license plate, traffic cones, high-visibility traffic safety vest and a traffic drum. The per-frame exposure time was set low to ensure points on the retroreflector remained under the count limit (4096 counts per timing bin per frame). Because of the ubiquity of roadsigns, our first experiments conducted were on a stop sign. The results of this test can be found in Fig. 4

Ground Truth For quantitative results, we collected the ground truth by masking the retroreflective elements with a combination of black construction paper and black tape. To our knowledge, no extensive glare datasets for LiDARs exist. One of the challenges to collecting such a dataset is acquiring ground truth data. One of the benefits of a non-learning based approach to glare correction is we do not need a large training set to achieve robust performance. Self-supervised approaches do not require ground truth data but they do require large training sets.

Comparison to Deep Learning Method Because there are limited academic works addressing glare suppression, we select a representative method [5] for comparison. [5] uses depth and intensity maps to model the GSF as an exponentially decaying function and uses synthetic data for training. The method suppresses glare artifacts in the depth map space. Because the training paradigm requires simple retroreflector geometry (e.g. octagons or rectangles) we observed this approach struggled on complex geometries such as the cone or high-vis vest. Scene independence is an important benefit of our method, which learning-based approaches lack without an extensive training dataset.

Aggregate Scene Results To demonstrate the generality and versatility of our approach, we show results on scenes with multiple retroreflectors of various types. For one scene we placed retroreflectors of varying size and shape at different depths. In another scene we emulated a scenario of a child crossing the road. For this we used a black mannequin which offered a challenging test case as weaker signals are more difficult to recover. We placed the mannequin at the same depth as the stop sign and a traffic drum to demonstrate that the signal can still be recovered when masked by bloom corruption in the same timing bins. We also show that retroreflector elements when worn on the person (in the form of a high-vis traffic safety vest) can still be suppressed while retaining the signal. This is an important component of our method because other glare suppression techniques are too aggressive and can remove the signal in addition to the corrupted region. This is problematic if a person is hid-

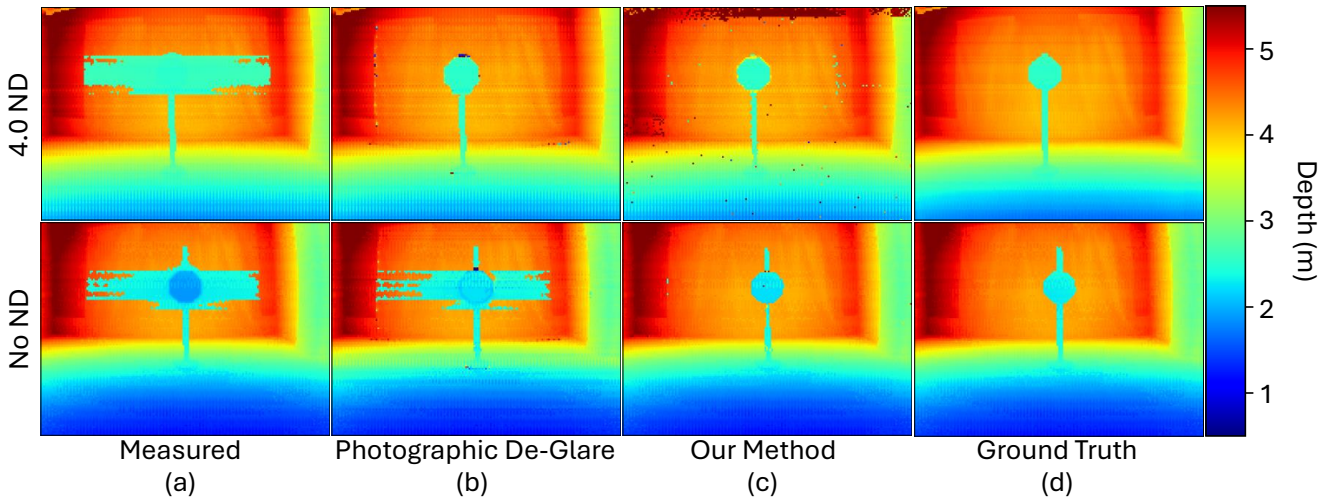


Figure 7. Comparison of our method to a photographic de-glare method that uses an algorithm proposed by Talvala et al. [25] to independently remove glare from each datacube time slice, as if it were a still image. (top) When pileup is mitigated by attenuating the signal with a neutral density filter (NDF, OD4), both methods mitigate glare successfully. (bottom) Without the NDF, severe pileup causes the photographic de-glare approach to fail.

den by the glare. The qualitative results of these evaluations are shown in Fig. 6 with quantitative results inset. For quantitative analysis we report δ_1 and RMSE. δ_i is given by:

$$\delta_i = \frac{1}{N}\sum_{j=1}^{N} \mathbf{1}(\max(\frac{d_j}{\hat{d}_j}, \frac{\hat{d}_j}{d_j}) < (1 + \frac{i}{100})) \qquad (9)$$

where $\mathbf{1}$ is the indicator function, N is the number of pixels, and $d_j, \hat{d}_j$ is the ground truth depth and predicted depth at the j-th pixel respectively. Because our method works in the echo space it can be easily integrated into existing LiDAR DSP pipelines.

Comparison to Photographic De-glaring Method Here we demonstrate how photon pileup causes photographic de-glaring methods to fail when applied to individual time slices of a raw SP-lidar datacube. We capture two sets of measurements for the scene containing a retroreflective stop sign. In the first, data is captured normally and histograms associated with pixels containing the stop sign show severe pileup distortion. In the second, a neutral density filter is placed in front of the lidar transmitter to reduce transmitted intensity by a factor of 10^4, which reduces the intensity of retroreflector returns such that pileup is negligible. The integration time of the second measurement is simultaneously increased (to 1.16 seconds), so that the total number of detected photons in both measurements are comparable.

For both sets of measurements, we treat each time slice of the captured datacube as an independent still image and attempt to remove glare from each time slice using a single-step glare removal operator[6] proposed by Talvala et al. [25]. We then employ a typical peak-finding procedure on the "de-glared" datacube, and form a depth map by selecting the depth of the brightest peak at each pixel. In Fig. 7 we compare depth maps obtained this way to depth maps produced with our method, which instead selects the depth of the highest confidence echo. Both methods successfully suppress glare when there is no pileup. However, in the case of the unmodified sensor, where pileup is severe, only our method, which operates on echos rather than the raw datacube, successfully suppresses the glare. The failure of the photographic de-glare method is due primarily to range-walk error induced by pileup. Pileup shifts direct returns from glare to earlier time bins, while the fainter glare signal is not shifted in time. This separates glare and aggressor signals into different time slices, such that linear de-glare methods operating on individual time slices over-predict glare in earlier time slices, and fail to remove the true glare.

[6] See Supplement for theoretical details.

5. Limitations and Future Outlook

One limitation of our work is that we have yet to test our method in outdoor traffic scenes where, in particular, solar background is more intense. Our method also requires pre-calibration of a GSF unique to each sensor and the validity of this GSF may degrade in realistic conditions—e.g. dust, condensation, or scratches on optical elements.

Looking forward, we expect de-glare to become a standard component of 3D perception pipelines—analogous to denoising or compression in conventional imaging. Historically, legacy LiDARs adopted suboptimal spinning architectures to avoid glare. Real-time algorithmic de-glare could reduce the need for such constraints, freeing designers to pursue fully solid-state, high-speed, and high-resolution architectures. The ability to correct glare algorithmically thus expands the design space for the next generation of "camera-fied" LiDARs.

References

[1] Waymo open dataset: An autonomous driving dataset, 2019. 2

[2] Brian F. Aull, Michael Collins, Kevan Donlan, Joseph Ciampi, Meera Punjiya, K. Alexander McIntosh, Kevin Ryu, and Deborah Woods. Mitigation of optical crosstalk in Geiger-mode avalanche photodiode arrays for lidar. In *Advanced Photon Counting Techniques XVI*, page 120890E. International Society for Optics and Photonics, SPIE, 2022. 3

[3] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. In *CVPR*, 2020. 2

[4] Edoardo Charbon, Matt Fishburn, Richard Walker, Robert K. Henderson, and Cristiano Niclass. *SPAD-Based Sensors*, pages 11–38. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. 1, 2

[5] Sungjin Cheong and Jusung Ha. Lidar blooming artifacts estimation method induced by retro-reflectance with synthetic data modeling and deep learning. In *2024 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, pages 1–4, 2024. 3, 6, 7

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

[7] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research (IJRR)*, 2013. 2

[8] Bhavya Goyal, Felipe Gutierrez-Barragan, Wei Lin, Andreas Velten, Yin Li, and Mohit Gupta. Robust 3d object detection using probabilistic point clouds from single-photon lidars. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 28417–28427, 2025. 3

[9] Anant Gupta, Atul Ingle, Andreas Velten, and Mohit Gupta. Photon-flooded single-photon 3d cameras. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6763–6772, 2019. 3

[10] Felipe Gutierrez-Barragan, Fangzhou Mu, Andrei Ardelean, Atul Ingle, Claudio Bruschini, Edoardo Charbon, Yin Li, Mohit Gupta, and Andreas Velten. Learned compressive representations for single-photon 3d imaging. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 10722–10732, 2023. 3

[11] Felix Heide, Steven Diamond, David B. Lindell, and Gordon Wetzstein. Sub-picosecond photon-efficient 3d imaging using single-photon sensors. *Scientific Reports*, 8(1), 2018. 3, 4, 5

[12] Atul Ingle and David Maier. Count-free single-photon 3d imaging with race logic. In *Proceedings of the 10th International Conference on Computational Photography (ICCP 2023)*. IEEE, 2023. 3

[13] International Organization for Standardization. ISO 9358:1994: Optics and optical instruments – Veiling glare of image forming systems – Definitions and methods of measurement, 1994. Standard. 3

[14] Jongho Lee, Atul Ingle, Jenu V Chacko, Kevin W Eliceiri, and Mohit Gupta. Caspi: collaborative photon processing for active single-photon imaging. *Nature Communications*, 14(1):3158, 2023. 2, 3

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

[16] Jiayong Peng, Zhiwei Xiong, Xin Huang, Zheng-Ping Li, Dong Liu, and Feihu Xu. Photon-efficient 3d imaging with a non-local neural network. In *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI*, page 225–241, Berlin, Heidelberg, 2020. Springer-Verlag. 3

[17] François Piron, Daniel Morrison, Mehmet Rasit Yuce, and Jean-Michel Redouté. A review of single-photon avalanche diode time-of-flight imaging sensor arrays. *IEEE Sensors Journal*, 21(11):12654–12666, 2021. 4

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

[19] Joshua Rapp, Yanting Ma, Robin M. A. Dawson, and Vivek K Goyal. High-flux single-photon lidar. *Optica*, 8 (1):30–39, 2021. 5

[20] Ivan Rech, Antonino Ingargiola, Roberto Spinelli, Ivan Labanca, Stefano Marangoni, Massimo Ghioni, and Sergio Cova. Optical crosstalk in single photon avalanche diode arrays: a new complete model. *Opt. Express*, 16(12):8381–8394, 2008. 3

[21] Mushfiqur Rouf, Rafał Mantiuk, Wolfgang Heidrich, Matthew Trentacoste, and Cheryl Lau. Glare encoding of high dynamic range images. In *CVPR 2011*, pages 289–296, 2011. 3

[22] Kaustubh Sadekar, David Maier, and Atul Ingle. Single-photon 3d imaging with equi-depth photon histograms. In *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part LXIV*, page 381–398, Berlin, Heidelberg, 2024. Springer-Verlag. 3

[23] Michael P. Sheehan, Julián Tachella, and Mike E. Davies. Spline sketches: An efficient approach for photon counting lidar. *IEEE Transactions on Computational Imaging*, 10: 863–875, 2024. 3, 5

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

[25] Eino-Ville Talvala, Andrew Adams, Mark Horowitz, and Marc Levoy. Veiling glare in high dynamic range imaging. *ACM Trans. Graph.*, 26(3):37–es, 2007. 1, 3, 4, 8

[26] Yoshiaki Tashiro and Kyosuke Ito. Spad depth sensor for automotive lidar systems. *JSAP Review*, page 230402, 2023. 3

[27] Satarupa Uttarkabat, Sarath Appukuttan, Kwanit Gupta, Satyajit Nayak, and Patitapaban Palo. Bloomnet: Perception of blooming effect in adas using synthetic lidar point cloud data. In *2024 IEEE Intelligent Vehicles Symposium (IV)*, pages 1886–1892, 2024. 3

[28] Max-Olivier Van Bastelaer, Heiner Kremer, Valentin Volchkov, Jean-Claude Passy, and Bernhard Schölkopf.

Glare removal for astronomical images with high local dynamic range. In *2023 IEEE International Conference on Computational Photography (ICCP)*, pages 1–11, 2023. 1, 3, 4

[29] Inc. Velodyne LIDAR. *HDL-64E S2 — User's Manual and Programming Guide*. Velodyne LIDAR, Inc., rev. h edition, 2012. Firmware version 4.07i. 2