

STORM: Super-resolving Transients by OverSampled Measurements

Ankit Raghuram^{*, 1}, Adithya Pediredla^{*, 1, 2},
Srinivasa G. Narasimhan², Ioannis Gkioulekas², and Ashok Veeraraghavan¹

^{*}Authors contributed equally

¹Department of Electrical and Computer Engineering, Rice University, USA

²Robotics Institute, Carnegie Mellon University, USA

Image sensors that can measure the time of travel of photons are gaining importance in a myriad of applications such as LIDAR, non-line of sight imaging, light-in-flight imaging, and imaging through scattering media. While the price of these sensors is dramatically shrinking, there remains a trade-off between spatial resolution and temporal resolution. While single-pixel detectors using the single photon avalanche diode (SPAD) technology can achieve 10-30 ps time resolution, the current generation array detectors can only produce an order of magnitude lower temporal resolution due to space-related fabrication constraints. Moreover, this limit is due to bandwidth, read-out and circuit-area constraints on the detector array and therefore unlikely to dramatically change in the next few years.

In this paper, we demonstrate a computational imaging approach that utilizes multiple measurements with calibrated sub-temporal resolution delays on the illumination pulse and super-resolution post-processing algorithms that together can achieve an order of magnitude improvement in the time resolution of the acquired transients. We build an experimental prototype, using a 32×32 SPAD detector array with 400ps time resolution and demonstrate recovery of transients with ≈ 50 ps time resolution, an $8\times$ improvement in time resolution resulting in a $5\times$ improvement in depth reconstruction error.

Index Terms—Computational Photography, SPAD camera, Temporal Super Resolution, Oversampling

I. INTRODUCTION

TRANSIENT IMAGING refers to imaging technologies allowing to record time-resolved photon flux, so-called transients, rather than steady-state intensity. Transient sensors are fast emerging as a critical technology in a variety of application domains, including 3D imaging [19], [31], [57], imaging through scattering [48], non-line-of-sight imaging [36], [54], autonomous navigation [11], and visualizing light-in-flight [55]. In the past two decades, a number of different methodologies for transient imaging have been developed, including interferometric techniques [14], streak cameras [55], single-photon avalanche diodes [13], ultrafast photodiodes [27], and continuous-wave time-of-flight (CW-ToF) sensors [20]. These methodologies provide different trade-offs in terms of temporal resolution (from a few femtoseconds for interferometry, to several nanoseconds for CW-ToF) and cost (from a few hundreds of dollars for CW-ToF sensors, to several hundred thousand dollars for streak cameras).

A. Single Photon Avalanche Diodes (SPAD)

Sensors based on SPAD technology offer a number of advantages compared to alternatives, making them well-suited for a majority of transient imaging applications: They provide sub-nanosecond temporal resolution. They can be fabricated using traditional CMOS fabrication processes enabling scalability. They can potentially be inexpensive, especially at volume given the scaling laws for CMOS fabrication [57]. They facilitate creation of detector arrays, allowing snapshot two-dimensional transient imaging to be realized in a compact

form-factor. These advantages make SPADs a promising and widely-applicable emerging technology for transient imaging.

B. Temporal Resolution Challenge

Single-pixel SPAD detectors have been demonstrated to achieve about 20 – 30 ps time resolution [10]. For applications like range imaging, this would correspond to 5 mm depth resolution. Unfortunately, two-dimensional arrays of SPAD detectors can only achieve an order of magnitude lower temporal resolution, due to space-related fabrication constraints. The two-dimensional SPAD detector array employed by us only provides a temporal resolution of approximately 400 ps [57], a $10 - 20\times$ lower time resolution. Again, for applications such as ranging this would yield a 5 – 10 cm depth resolution, something that is not suitable for most emerging applications.

The main challenge in fabricating two-dimensional arrays of SPAD detectors lies in the circuitry required per-pixel to obtain photon timing information. This is typically achieved through time-to-digital converter (TDC) circuits. The timing jitter that can be realized on these circuits depends on the area available to implement them. This sets up a tradeoff between light efficiency and temporal resolution, as the same physical area cannot be used for light detection and TDC circuits. As many applications utilizing transient imaging are light-starved, this tradeoff poses a formidable challenge. In current arrays [34], the pixel active area is less than 10% of the total area, and even with this, the temporal resolution that can be achieved is limited to several hundreds of picoseconds. Therefore, there is strong demand for *computational imaging* techniques that can help overcome this trade-off, at least until advances in fabrication technology can provide the order of magnitude improvement in temporal resolution needed for SPAD array detectors, without incurring a severe loss of light efficiency.

C. Key Idea: Time-shifted Measurements

The key insight in this paper is that temporal resolution in 2D SPAD arrays is lost in the time-to-digital conversion process—that is, in a process that is analogous to the box filtering implemented by pixels and impacting spatial resolution in traditional intensity sensors. It is well-established in quantization theory [5] that sub-quantization level recovery is possible if multiple shifted measurements are obtained. We leverage this same principle in the context of the time-to-digital conversion to improve temporal resolution.

In particular, we propose to make multiple measurements with slightly phase shifted (time-shifted) illumination pulses – wherein the time shift is an order of magnitude smaller than the temporal resolution. This amounts to shifting the quantizer levels by less than the quantizer step-size. These additional measurements ensure that sufficient information is available in the captured data to reconstruct the signal with much higher temporal resolution than that afforded by the on-chip TDC circuits. We call our technique *STORM: Super-resolving Transients by OverSampled Measurements*.

We demonstrate the effectiveness of our technique using a commercially-available SPAD array sensor. This sensor, which is representative of the current limits of 2D SPAD array technology, has total timing jitter on the order of a few tens of picoseconds, and TDC resolution of about 400 picoseconds. Therefore, it is well-suited for our technique, which assumes that TDC resolution is the main bottleneck limiting temporal resolution. Indeed, in our experiments, we show that using STORM can result in around $8\times$ improvement in temporal resolution. Even though our focus is on SPADs, STORM can improve the temporal resolution of other transient detectors such as streak cameras, lidar, ICCDs¹.

D. Contributions

We summarize our technical contributions:

- 1) We propose STORM, a transient super-resolution technique that uses multiple time-shifted measurements to increase the resolution of transient detectors that are limited by the TDC quantization step-size.
- 2) We exploit well-known results in quantization theory to mathematically model the measurement process of a 2D transient detector array. We derive from this model a deconvolution technique that uses multiple time-shifted measurements to obtain higher temporal resolution.
- 3) We use simulated data and measurements from an experimental SPAD setup to demonstrate that STORM can increase temporal resolution by a factor of $30\times$ in simulation, and up to $8\times$ in experiments. This in turn has improved depth reconstruction error by $15\times$ in simulations and $5\times$ in experiments.

E. Limitations

STORM acquires multiple measurements through time-multiplexing: successive captures at different time shifts. This

can result in increased total capture-time, though this is not always the case. As we show in our experiments, the multiple measurements can often be obtained using the same *total* exposure time, with each individual measurement having a lower exposure time and therefore signal-to-noise ratio.

STORM relies on the observation that, in current SPAD arrays, the TDC circuits are the main bottleneck limiting temporal resolution, with both the timing jitter and the illumination pulse-width being much lower. In situations where either the timing jitter or the illumination pulse-width are the primary cause of lower temporal resolution, STORM will likely not be effective (see Section IV-B for some quantitative discussion). Examples of this include single-pixel SPAD detectors and 1D line SPAD sensors: In these sensors, there is sufficient real estate around each pixel to realize higher temporal resolution TDC circuits. Another example is when the application permits the use of off-chip samplers providing very high-resolution sampling (e.g., PicoHarp at 4 ps or HydraHarp at 1 ps [43], [44]).

II. RELATED WORK

Imaging beyond the resolution limits imposed by hardware (e.g., lenses, sensors) and physics (e.g., diffraction) is a long-standing challenge in all aspects of imaging, and has led to several scientific discoveries and commercial products. Below, we review some of the spatial and temporal super-resolution techniques that are related to the ideas presented in this paper.

A. Super-resolution Microscopy

Due to the wave-nature of light, diffraction causes image blur and limits the maximum spatial resolution of light-based imaging systems. This limit is given by Abbe's limit: $\frac{\lambda}{2NA}$, where λ is the wavelength of light wave and NA is the numerical aperture of the imaging system. This has long-motivated research for overcoming the *diffraction limit* in spatial resolution. A well-established approach is to exploit fluorescence, and the development of fluorescence-based super-resolved microscopy was recognized by the 2014 Nobel Prize in Chemistry [8]. For instance, in stimulated emission depletion microscopy (STED) [21], an excitation laser is used to create a diffraction spot, and a depletion laser, in the shape of a doughnut, de-excites the fluorophores around the excitation laser spot, creating a sub-diffraction laser spot. In structured illumination microscopy [3], [17], [18], the projection of sinusoidal patterns is used to alias high-frequency components, which are captured as low frequency components that are within the frequency limit imposed by the diffraction blur kernel. In photo-activated localization microscopy [8] and stochastic optical reconstruction microscopy [47], sequential activation and time-resolved localization of photo-switchable fluorophores create sparsity in the images. Exploiting this sparsity structure, these techniques localize the fluorophores accurately, leading to image super-resolution.

A different methodology for spatial super-resolution, and one closer to this work, is Fourier ptychography [53], [60]. In this case, the diffraction limit imposed by lenses is overcome by translating either the lens or the light source, in order to

¹Concurrent to this work, Cester et al. [9] proposed temporal super resolution for Intensified CCDs with ideas similar to STORM.

create a synthetic numerical aperture larger than the numerical aperture of the imaging system. This approach also finds applications in macroscopic imaging, where size and cost make lenses with large apertures impractical [23]. Analogously to how ptychography techniques sweep the Fourier plane through calibrated shifts of the lens, in STORM we sweep the temporal axis by time-shifting the laser pulse signal.

B. Video super-resolution

Geometric super-resolution techniques exploit sub-pixel camera motion between successive video frames in order to increase the video's spatial resolution. Peleg et al. [41] first proposed exploiting camera motion in a video to achieve spatial super-resolution. Several subsequent techniques [33] improved upon this initial approach, by using better image registration [50], [59] or blur estimation and deblurring algorithms [2], [28], [49].

Ben et al. [6], [7] introduced the so-called *jitter camera*, where a micro-actuator is used to control the exact location of the sensor behind the lens. By shifting the sensor at sub-pixel increments during video capture, this enables recovering video at higher spatial resolutions. Yu et al. [58] adapted this technique to optical microscopy, using micro-mirrors to shift the location of the image on the sensor.

Our proposed transient super-resolution technique is conceptually similar to the above-described approaches for geometric video super resolution: Instead of spatially shifting an image by sub-pixel amounts, we temporally shift a transient by sub-bin-width amounts. Additionally, Baker et al. [4] and Lin et al. [30] computed the fundamental limits for spatial superresolution techniques. We derive similar limits for temporal superresolution.

C. High resolution transient cameras

Transient imaging cameras can be broadly separated into two categories, based on the underlying imaging principle [25]. First are *impulse-response* transient cameras: These measure transients by probing the scene with a pulse of light of narrow temporal width, and capturing time-resolved measurements of the light returned from the scene. This category includes streak cameras (resolution 2 ps) [56], SPAD cameras (30-400 ps) [13], [35], [57], ultrafast photodetectors (30-400 ps) [27], and intensified CCD (ICCD) sensors (200 ps) [24]. Second are *correlation* cameras, which reconstruct transients by measuring their correlation with continuous-time waveforms. These include continuous-wave amplitude-modulated time-of-flight sensors, which can achieve temporal resolutions ranging from 10 ps to 1 ns [20], [26], [42], [52]. This category additionally includes interferometric techniques such as optical coherence tomography (OCT), which can achieve resolutions of 33 fs [14]. Despite the availability of transient imaging technologies with higher temporal resolution than SPADs, most of these alternatives are not realistic options when it comes to deploying at scale or in small form-factors. Therefore, developing computational imaging techniques for increasing the temporal resolution of SPADs remains an important research objective.

Recently, a number of techniques have been used to computationally increase temporal resolution of SPADs to subpicosecond levels, for the specific application of depth ranging (i.e., transients consisting of a single return) [19], [45]. Compared to these works, our technique is a general-purpose super-resolution procedure that can be employed for arbitrary transient signals and applications of transient imaging. Gruber et al. [16] showed that an accurate depth map can be captured with 3 time-gated images using a neural network architecture. Finally, Lindell et al. [31] proposed increasing the spatial, rather than temporal, resolution of SPAD cameras for depth sensing applications, by combining SPAD measurements with high-resolution measurements from an intensity camera.

III. TRANSIENT IMAGING WITH SPADS

In this section, we briefly describe transient acquisition with SPADs, and provide a mathematical model for the acquisition process, which we will use to derive the STORM reconstruction procedure. We note that, as STORM is independently applied for each pixel of the SPAD array, we focus only in acquisition at the level of individual pixels. We additionally assume that each pixel in the SPAD array operates independently. This is a realistic assumption, only failing to account for a small cross-talk effect between the neighboring pixels [57].

We emphasize that, due to our focus on light-starved applications, both our high-level description and mathematical model assume that the SPAD is operating in the so-called *low flux* regime (see, for instance, Shin [51] for a thorough discussion). In the presence of strong incident flux on the SPAD, this assumption is not valid and it becomes necessary to consider SPAD operation in the *high flux* regime [19], [39], [46].

We begin by providing a high-level overview of SPAD operation, deferring to Villa et al. [57] for details. To measure the transient response of a scene, SPADs are typically coupled with a pulsed illumination source. The source emits an ultra-short pulse towards the scene, where photons scatter, and some of them return to the SPAD pixel. Measuring the true transient response of the scene would require ideally binning *all* of the returning photons based on their time-of-travel. However, SPADs detect at most one among the returning photons that originated in the same pulse (assuming operation in the low-flux regime). Therefore, in order to form a full transient, the above process is repeated several times, by successively emitting laser pulses and detecting and binning one returning photon at a time.

Detected photons are binned based on their time-of-travel using a time-to-digital converter (TDC) circuit [57]. Multiple TDC techniques such as Vernier delay line, flash TDC, time-correlation single photon counting (TCSPC) are used in designing TDCs [1]. The detector we have employed uses the TCSPC technique, which we explain next. For implementing TCSPC, the TDC contains a capacitor that is charged with a constant voltage during the travel duration of the photon (time between the emission of the photon packet from the laser and the detection a single photon by the SPAD). Therefore, ideally the final charge is proportional to the total time-of-travel of

the detected photon. In practice, this is not the case exactly, because of the finite temporal width of laser pulses (photons in the same pulse have different emission times, but capacitor charging starts at the same time) and random errors in timing photon detection events, often referred to as the timing *jitter*.

The TDC also contains an analog-to-digital converter (ADC) circuit, which converts the analog charge to a digital value. By forming a histogram of the detected photons based on this digital time-of-travel, this process eventually creates an estimate of the transient response of the scene.

The least significant bit (LSB) of the ADC determines the temporal resolution of the transients. For the SPAD camera we used in our experiments, the LSB results in a temporal resolution of around 400 ps (the exact value reported in the specifications is 401.875 ps). By comparison, the laser pulsewidth and timing jitter are in the order of 10–20 ps each. Therefore, the finite resolution of this ADC is the main limiting factor when it comes to the ultimate temporal resolution of the transient imaging system. Before introducing our methodology for overcoming this limit, we provide a mathematical model describing this transient acquisition process.

A. Mathematical formulation

Let $x(t)$, $t \in \mathbb{R}_+$ be the ideal transient response of the scene we are imaging. Let $h(t)$ be the total system jitter, which includes the laser pulse jitter and also the SPAD jitter. Finally, let $B_T(t)$ be the box function with a width T equal to the temporal resolution of the TDC. Then, under the low-flux model [51], and as the number of emitted pulses approaches infinity, entries of the histogram measured by the SPAD can be modeled as independent Poisson random variables,

$$y(n) \sim \text{Poisson} \left(c \cdot \int_0^{T_{\max}} (x \otimes h \otimes B_T)(t) \delta(t - nT) dt + \eta \right), \quad (1)$$

where $\otimes$ denotes circular convolution, $n \in \{0, 1, 2, \dots, N\}$, c is a scalar depending on the photon detection efficiency of the SPAD and the number of emitted pulses, and η corresponds to photon detections due to dead counts. N and $T_{\max}$ typically depend on the interval between successive laser pulses (the so-called *repetition rate*), with $N = \frac{T_{\max}}{T}$ typically in the order of tens of thousands. We note that it is possible for the scene transient $x(t)$ to extend beyond $T_{\max}$, in which case we need to instead consider the *wrapped transient*

$$\tilde{x}(t) = \sum_{l=0}^{\infty} x(t + l \cdot T_{\max}), \quad t \in [0, T_{\max}]. \quad (2)$$

For simplicity, in the rest of the paper, we will not discriminate between $x(t)$ and $\tilde{x}(t)$. Additionally, for ease of notation, we will drop the integration limits and the terms c and η from Equation (1).

From Equation (1), we note that, unless the scene response $x(t)$ (or $x \otimes h$) is band limited with maximum frequency $\frac{2\pi}{T}$, transient events that occur within the same sampling period T are aliased and lost irrecoverably (unless we have scene priors). In the next section, we propose STORM that changes the transient acquisition process to preserve the details of transient events that happen within each sampling period T .

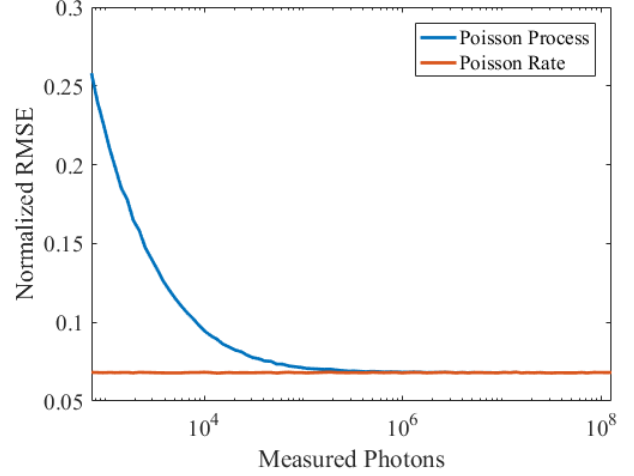


Fig. 1: **Poisson process dominates only for low photon count.** We reconstruct a ground-truth, noiseless, high-resolution transient, by applying STORM to two types of lower-resolution transient signals (Equation (4)): The Poisson rate $r(k)$ (red curve), and Poisson processes $y(k)$ (blue curve) corresponding to exposures of increasing duration (increasing photon counts). We observe that the addition of Poisson noise has a negligible effect on the reconstruction error as the photon count increases.

IV. DESCRIPTION OF STORM TECHNIQUE

To acquire transients that retain frequency content higher than the Nyquist limit imposed by the sampling bin width T , we propose to take multiple measurements by time-shifting the scene transient $x(t)$. For STORM to be effective, the time-shifting step needs to be smaller than the temporal resolution². The time shift can be realized by either delaying the laser beam using an optical delay line (common practice in optical coherence tomography), or by delaying the laser synchronization signal sent to the TDCs. In our experiments, we adopt the second approach, and implement the delay using a picosecond delay (see Section VI for implementation details).

The transients measured with the delay are given by

$$y_m(n) \sim \text{Poisson} \left(\int (x \otimes h \otimes B_T)(t) \delta(t - mS - nT) dt \right), \quad (3)$$

where S is the time-shifting step, and $m \in \{0, 1, 2, \dots, \frac{T}{S} - 1\}$.

Equation (3) can be further simplified as

$$y(k) \sim \text{Poisson} \left(\underbrace{\int (x \otimes h \otimes B_T)(t) \delta(t - kS) dt}_{r(k)} \right), \quad (4)$$

where $k \in \{0, 1, 2, \dots, N \frac{T}{S}\}$. This is similar to Equation (1), with an oversampling factor of $\frac{T}{S}$. Therefore, the proposed acquisition technique STORM samples the *box-and-jitter filtered* transients at a higher temporal resolution.

²In this manuscript, we focus on fixed time-shifts. However, temporal superresolution can also be achieved using random time-shifts, analogous to the spatial superresolution case [30]. We leave this for future work.

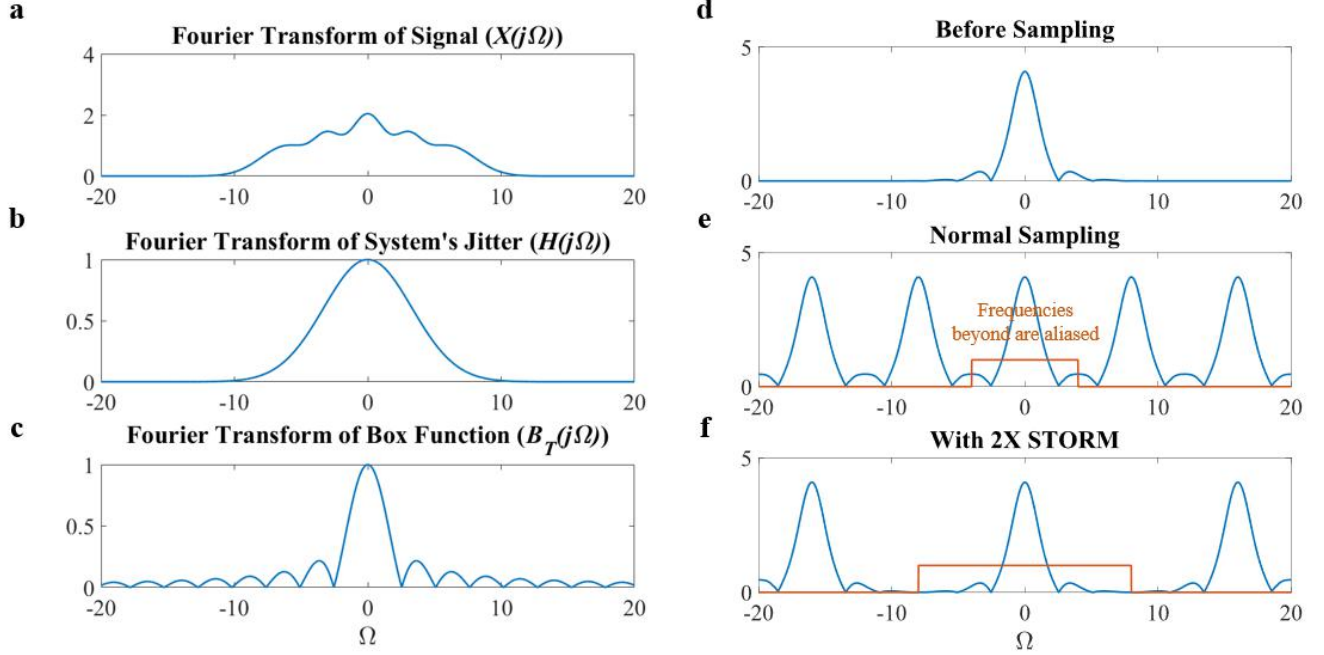


Fig. 2: **Fourier interpretation of STORM:** **a** The Fourier transform of a typical transient is not band limited due to sharp peaks in the time domain. **b** The system jitter can be approximated with a Gaussian shape. **c** The Fourier transform of box function is a sinc function. Note that sinc causes zeros at all integer multiples of the sampling frequencies and these frequencies cannot be recovered without any prior knowledge. **d** Fourier transform before sampling. **e** Sampling introduces periodicity and aliasing causing loss of information **f** STORM prevents aliasing. Note that frequency content lost due to system jitter cannot be recovered by STORM.

A. Fourier-domain interpretation

The effect of various parameters (system jitter h , temporal bin width T , time shift S) can be understood elegantly in the Fourier domain. To simplify analysis, instead of the Poisson process in Equation (4), we work directly with the rate term $r(k)$. Figure 1 provides some justification for this simplification, by showing that the additional reconstruction error from ignoring the Poisson noise becomes negligible for sufficiently large exposures. The Fourier transform of $r(k)$ is

$$R(j\Omega) = \sum_{k=-\infty}^{\infty} (X \cdot H \cdot B_T)(j(\Omega - k\Omega_S)), \quad (5)$$

where $\Omega_S = \frac{2\pi}{S}$. Figure 2 shows the magnitude of the various terms in Equation (5). The system jitter h is approximated with a Gaussian function. This is because transform-limited laser pulses are Gaussian in shape, and because SPAD jitter is well-modeled as the combination of a Gaussian and a heavy-tailed exponential distribution, with the exponential tail being two orders of magnitude smaller than the Gaussian [22].

Without STORM ($S = T$), aliasing destroys the high-frequency content of the transients and, worse even, corrupts the low frequencies. Employing STORM with an oversampling factor of $2\times$ ($S = \frac{T}{2}$) decreases aliasing. In an ideal setting (no noise or system jitter), we can use arbitrarily large oversampling factors and continue improving the temporal resolution.

In practice, noise and system jitter place limits on how much we can super-resolve, as we discuss next.

B. When to use STORM

From Figure 2, we notice that if frequency content is lost due to the system jitter, STORM cannot retrieve the lost frequencies. In the presence of noise, the frequencies that are below the noise floor cannot be recovered either. Therefore STORM should be employed for SPAD cameras that have time bins in the order of, or larger than the system jitter. Formally, if the width (corresponding to highest non-zero frequency) of the system jitter is τ , we can express the limits of STORM as:

$$\text{optimal super-resolution factor } \frac{T}{\tau} \text{ with shift } S = \tau. \quad (6)$$

Experimentally, we have observed improvements with oversampling factors as high as $40\times$ (see Section VI).

C. Reconstructing with STORM

Even though STORM increases the sampling resolution, the resulting measurements are still blurred by the system's jitter and the box function, as well as corrupted by Poisson noise. Reconstructing the transient x in Equation (4) requires solving a simultaneous denoising and deconvolution problem. We approximately solve this problem using Wiener

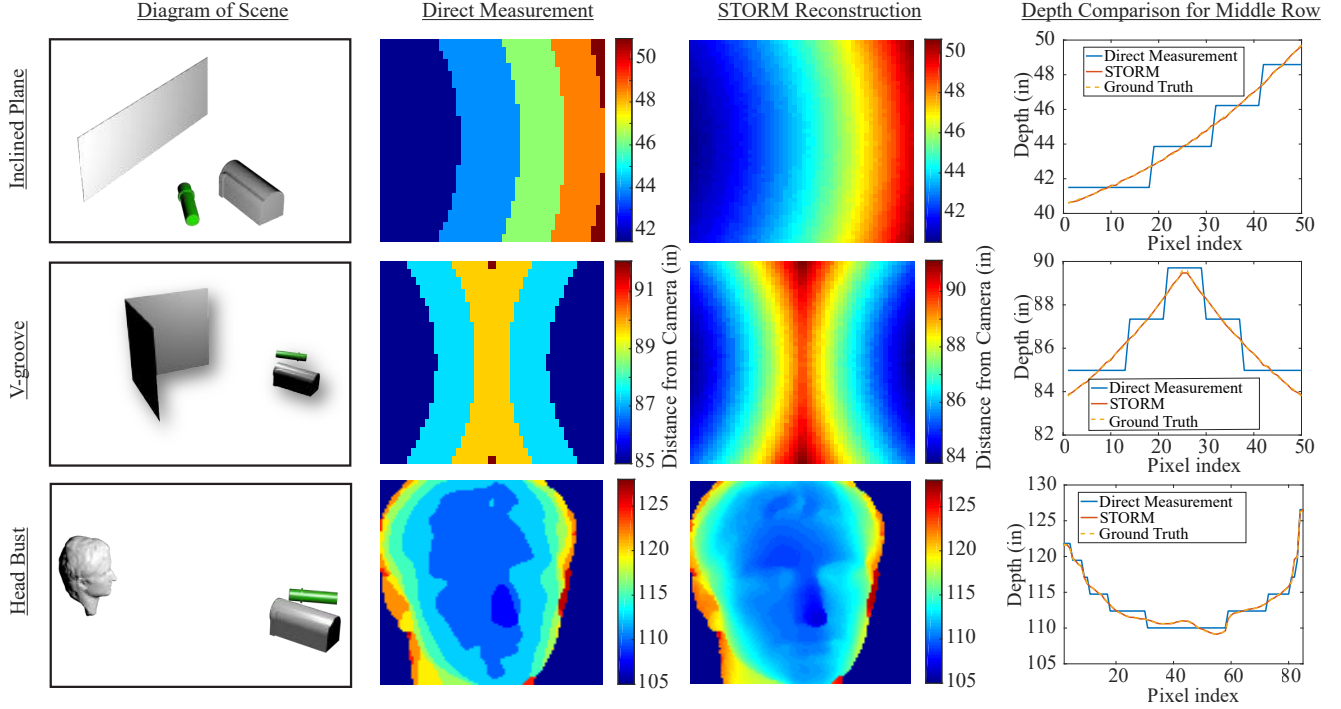


Fig. 3: **Simulated results on depth reconstruction:** Three scenes are used to investigate the impact of STORM on depth imaging: an inclined plane, a V-groove, and a human head bust. The left column shows the scene set up. The second and third columns show the computed depth map using direct measurements and STORM reconstruction. The fourth column shows a depth comparison of the two techniques along with the ground truth. Since the bust did not fill the entire image, locations not corresponding to the bust were given the lowest value for contrast. RMSE in inches for (direct, STORM) are (0.4531, 0.0326), (0.7037, 0.043396), and (0.6644 0.0426) for the inclined wall, V-groove, and bust, respectively. We observe around $15\times$ depth error reduction and a $30\times$ increase in temporal resolution with STORM.

deconvolution: Even though Wiener deconvolution is derived under an assumption of additive noise, which is not a valid assumption for the Poisson noise of Equation (4), in practice we found that it performs satisfactorily when applied to SPAD measurements. Our Wiener filter is

$$W(j\Omega) = \frac{B_T^*(j\Omega)}{|B_T(j\Omega)|^2 + \alpha(j\Omega)}, \quad (7)$$

where α is an SNR dependent noise term. In this work, α is chosen empirically. We exclude the system jitter h since its shape and extent are not known without extensive calibration. Thus, the effects of this jitter will still be present in our reconstructions.

We note that, when reconstructing “delta-like” transients x , Wiener deconvolution may also introduce ringing artifacts. This is especially detrimental when reconstructing transients with multiple features. To address this, in applications where reconstructing the entire transient beyond the first return is critical (e.g., Sections V-B and V-C), we reconstruct the transient x by using TVAL3 [29] to solve a TV-regularized least-squares problem. We use the authors’ implementation with default parameters, except for the primary penalty term (opts.mu), which we set to 2^4 . We also enforce that the signal is positive and satisfies an L_2 prior (smooth).

V. SIMULATIONS

In this section, we systematically evaluate STORM’s performance on scenes with increasing complexity simulated with a physically-based, open-source transient renderer [37], [40]. We quantify the accuracy of the signals reconstructed using STORM by comparing with ground-truth high-resolution renderings. The simulated imaging system has temporal resolution $T = 400$ ps. In addition to rendering noise, Poisson noise and picosecond delay jitter (5 ps [32]) are included in rendered transients. STORM is implemented using temporal shifts of step $S = 10$ ps, corresponding to an oversampling factor of 40. The *total* exposure duration is kept constant between a single direct measurement and the multiple measurements of STORM by scaling the intensities of low resolution renderings by the oversampling factor.

A. Depth Imaging

Measuring scene depth finds several applications such as automobile navigation (Lidar), simultaneous localization and mapping (SLAM), autonomous drones, and entertainment (XBox). For depth extraction, we measured the time-of-travel corresponding to the maximum of the transient. A system with higher temporal resolution will accurately estimate depth. The depth resolution can be approximated as $r_d = c \frac{T_t}{2}$ where r_t is the temporal resolution of the imaging system and c

is the speed of light. The depths of scene points that are closer to each other than r_d cannot be distinguished and lead to inaccurate scene reconstruction. To quantify the depth resolution of STORM we considered three scenes: (i) an inclined plane; (ii) a V-groove; and (iii) a human head bust.

An inclined plane (such as a ramp) varies in depth along one or two spatial dimensions (in our case only one). Since it is a fixed plane, only single bounce photons will arrive at the detector, resulting in a single sparse transient response smoothed by the total system jitter, $h(t)$. A schematic of the scene is shown in the first row of Figure 3. The center of the plane is 45 inches away from the detector and makes a 63 deg angle with the camera axis. The second and third columns of Figure 3 display reconstructed depth maps for the scene. When reconstructing depth directly from the low-resolution measurements (“direct measurement”), the depth map appears quantized into five distinct depth levels: This is because the total depth range of this scene is approximately 12 inches, which corresponds to 2 ns travel time, or five temporal bins in our simulated imaging system. By contrast, the depth map using the STORM technique has smoothly varying depths. There are approximately 150 distinct depth levels across the image in this case, indicating a $30\times$ improvement in temporal resolution using STORM. The last column of Figure 3 displays the depth estimate of the middle row of the sensor. Direct measurement of the scene manifests as discrete steps, making it ambiguous what the depth is between the depth levels (RMSE = 0.6644 inches). The reconstructed depths using STORM look continuous and align well with the ground truth (RMSE = 0.0426 inches) with an improvement over direct measurement of approximately $16\times$ in depth estimation. We note that the depths for the inclined plane show as a curved line which is a consequence of our use of a *perspective* camera: light from scene points farther away from the camera axis have to travel farther than light from scene points closer to the camera axis. This is also seen in the depth maps for the inclined plane, where the depth isocontours are curved.

Unlike the inclined plane, the V-groove scene (Figure 3, 2nd row) allows for multi-bounce photons to arrive at the detector, manifesting as a tail in the transient response. This can introduce bias into the depth reconstruction algorithms that use data priors. STORM is independent of data priors and does not suffer from this bias. The center of the simulated V-groove is located 90 inches from the detector and is composed of two planes that form a 130 deg angle. The depth maps for V-groove show discretization properties similar to the inclined plane. As evident in the depth comparison plot, the corner cannot be resolved with direct measurements (RMSE = 0.7037 inches) but can be resolved with STORM (RMSE = 0.0426 inches). The error reduction is approximately $16\times$.

The bust scene (Figure 3, 3rd row) introduces detailed depth features (indents from eyes, protrusion from nose, etc.) that are not present in the other two scenes. With STORM, many of the depth features are preserved, while with direct measurements, the features fall within the same depth level and are lost. One example can be seen in the bust scene comparison plot in the bottom-right part of Figure 3: The changes in depth between pixels 40 and 60 for the ground truth

and STORM reconstruction correspond to the depth profile of the nose. This profile is lost for direct measurement. Direct measurement of the middle row of the image resulted in an RMSE of 0.4531 inches, and STORM resulted in an RMSE of 0.0326 inches. This is an improvement of approximately $12\times$.

B. Non-Line of Sight (NLOS) imaging

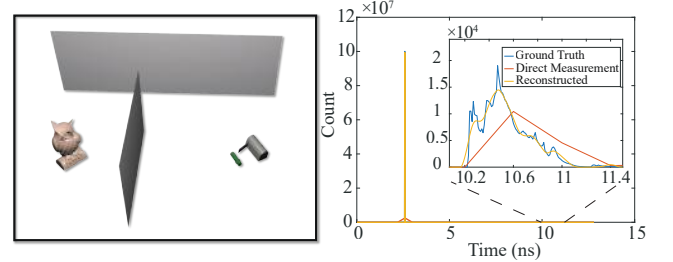


Fig. 4: STORM enables higher fidelity third-bounce transient capture for looking-around-the-corner applications. As shown in the diagram (left), the owl is not in the line of sight of the source and detector. The plot on the right shows the transient response of the scene measured directly and using STORM. The inset shows the time interval 10 – 11.5 ns, to highlight the part of the transient due to third-bounce photons coming from owl. STORM accurately traces the profile of the third-bounce photons, which is lost in direct measurement.

The STORM technique allows us to recover super-resolved versions of the entire transient, and not just the first return required for depth sensing. Reconstruction of the full transient allows inspection of later bounces (e.g., third-bounce photons), which is useful in applications such as non-line-of-sight imaging [36], [38], [54]. NLOS algorithms use third-bounce photons to reconstruct objects that are not directly in the field of view of the camera. In our example (Figure 4), we place an owl bust at a location that is not directly visible by the detector and source. The detector and source are co-located and are focused at a point on the visible wall (similar to the hardware geometry used by O’Toole et al. [36]). The corresponding transient is shown in Figure 4 on the right. (Note that the transient is dominated by the single-bounce peak, so we zoom in to highlight the part of the transient due to third-bounce photons.) Without STORM, most of the third-bounce transient is lost. With STORM, both the single-bounce peak and the third-bounce profile are recovered accurately.

C. Volumetric scattering media

Satat et al. [48] demonstrated that time-resolved measurements can be used to benefit applications attempting to image through thick scattering media. Temporal resolution is especially important when performing time-gating to separate out ballistic photons. The better the temporal resolution, the smaller the amount of scattered photons arriving within the same time window as the ballistic photons. In Figure 5, we show that using this type of ballistic imaging with time-gated measurements, we can get a clear picture of the “ICCP” text that is occluded from the sensor by a scattering volume.

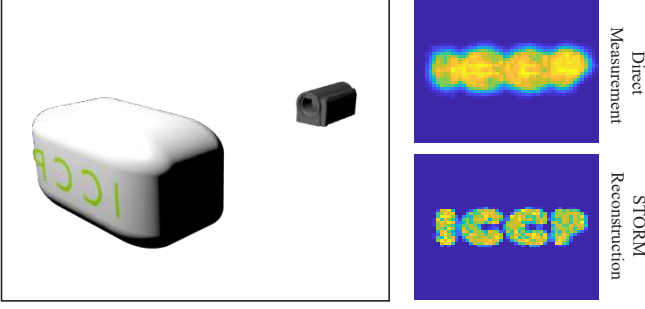


Fig. 5: **Higher resolution STORM transients enable better time-gating for imaging through scattering media:** Simulation of ballistic imaging of “ICCP” text hidden behind a scattering volume. The text emits light which travels through the scattering volume (approximately 9 mean free paths thick), and is imaged by an orthographic camera (set up similar to Satat et al. [48]). The images show the first frame where ballistic photons arrive. Acquisition of higher resolution transient using STORM allows a more accurate reconstruction of the text profile.

The volume has scattering coefficient $\sigma_t = 19.2 \text{ cm}^{-1}$ and geometric thickness 1.73 m, which corresponds to $9 \times$ the mean free path. The scattering phase function follows the Henyey-Greenstein parametric form [15] with an anisotropy parameter $g = 0.9$. We use an orthographic camera for this simulation. At the native temporal resolution of the transient detector, the frame that contains the ballistic photons also has strong contributions from multiply-scattered photons, resulting in strong blurring of the text. Without prior information, the image formed using direct measurements is difficult to read. With STORM, we see a sharper image, as there are fewer scattered photons being imaged. Both images correspond to the first photon frame of their respective transients.

VI. EXPERIMENTAL RESULTS

The following experimental results are obtained using a physical implementation of a SPAD-based transient imaging setup. We use a pulsed laser (NKT Photonics EXW-12) with a pulse-width of about 10 ps and pulse repetition rate of 77.8 MHz, as well as a 32×32 SPAD array (MPD SPC2 [12]) with bin-width $T = 401.375$ ps and internal clock at 155.1 MHz. Light is focused onto the SPAD with the help of a 12 mm c-mount lens (Computar M1214-MP2). Each pixel of the SPAD camera has an independent TDC circuit, which receives a sync signal from the laser and photon-triggered signal from the SPAD. To create the time-shifts required for STORM, instead of directly sending the laser sync signal to the TDCs, we used a picosecond delay (PicoQuant PSD) to programmatically delay the sync signal, and hence the transient. The PSD has a temporal resolution of $S = 10$ ps. For wide-field experiments, the laser beam is expanded with the help of a concave lens (25 mm diameter and 25 mm focal length, from Edmund Optics), which creates diffuse illumination that lights up the entire scene. For depth scenes, ground truth measurements are hand-measured and are therefore not

very accurate, and intended for qualitative comparisons. Furthermore, we ignore the effect of perspective distortion. We maintain the exposure duration constant between $k \times$ STORM and direct measurements, by keeping the integration time for each shifted measurement to be $\frac{1}{k} T_{acq}$, where T_{acq} is acquisition time for the direct measurements.

A. Depth Imaging

Similar to the simulations of Section V, we used an inclined plane to conduct an experiment involving a single-bounce scene. A white cardboard plane is kept 50 inches away from the detector at an angle of 64 deg relative to the camera’s optical axis (Figure 6). The depth maps in the second and third columns of Figure 6 show agreement with the inclined plane simulations (Figure 3). The direct measurement depth map displays approximately five discrete depth isocontours (two of them are visible only partially), whereas STORM reconstruction creates a smoother depth profile, increasing the number of discrete depth levels by a factor of 30 (170 levels). Moreover, when comparing with the ground truth, STORM improves RMSE by a factor of 3.3. As in the simulations, STORM reconstruction agrees well with the ground truth, whereas direct measurements bin the depths into discrete “steps”.

We perform a second depth sensing experiment for a V-groove scene, created by placing two cardboard planes at an angle of 140° relative to each other, as shown in Figure 6. As seen in the figure, reconstruction with STORM clearly recovers the “V” shape of the scene. On the other hand, direct measurement does not reliably recover the shape of the scene and is unable to reconstruct the corner formed by the “V”. Comparing recovered depth to the ground truth, we observe that STORM improves RMSE by about $5 \times$. We believe the difference in performance for simulated and experimental situations is largely due to model mismatch. In the simulation, the SPADs jitter is assumed to be a delta function. However, in reality, this will be a linear combination of a Gaussian with an exponential tail. The SPAD jitter can be incorporated in the forward model. However, with the current camera, we cannot measure this SPADs jitter as it is smaller than the bin-width of the system.

Before concluding this section, we briefly discuss how we determined the ground truth depths. Since both scenes are made of planes, the ground truth depths are computed using the distance of the plane from the camera (along the camera axis) and the angles of the planes relative to the camera axis. For simplicity, we start by computing the ground truth for a single inclined plane. The distance traversed by photons from the source-detector pair to a point on the plane is given by $d_t = \frac{d_z}{\cos \theta}$ where d_z is the depth of the plane (along the camera axis) and θ is the angle between the camera axis and the line starting from the source and ending at the point on the plane. If d_o is a point on the plane that intersects the camera axis ($\theta = 0$), then $d_z = d_o + x \tan \phi$ where $x = d_z \tan \theta$ and ϕ is the angle made between the plane and the imaging plane. Since $\tan \theta = \frac{x}{d_z}$, the final expression used to compute d_t is

$$d_t = \frac{d_o + x \tan \phi}{\cos(\tan^{-1}(\frac{x}{d_z}))}, \quad (8)$$

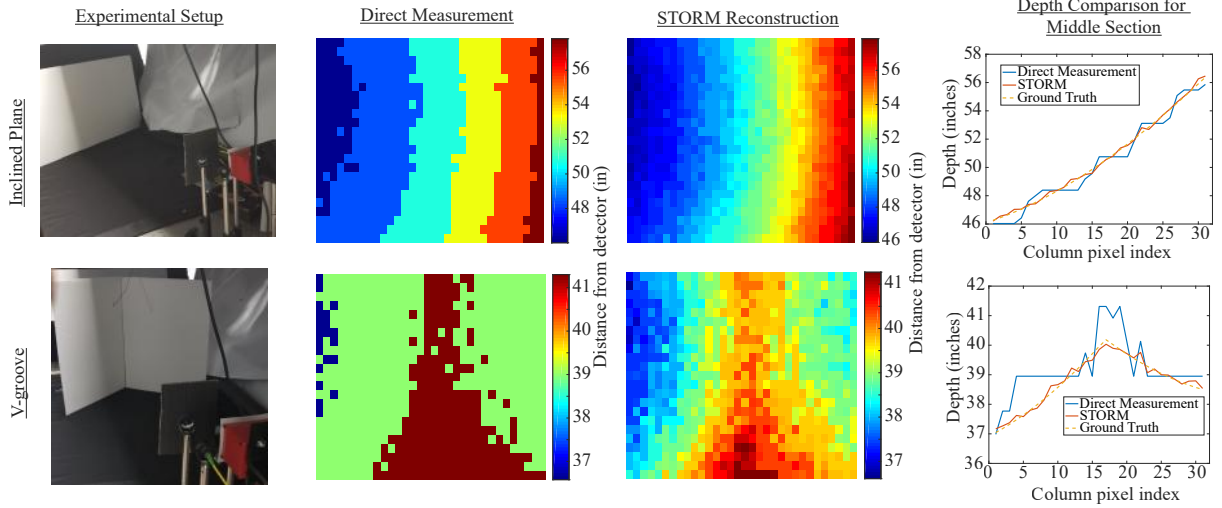


Fig. 6: **Experimental depth imaging using STORM.** The two rows of the figure correspond to the two depth imaging scenes: the inclined plane and V-groove scenes. The first column displays the experimental setup. The second and third columns show depth maps of the direct measurement and STORM techniques. The fourth column shows a depth comparison of the two techniques with the ground truth for an average of 5 rows. Ground truth estimate is shown below in eq. 8. RMSE in inches (direct, STORM) are (0.4855, 0.1427) and (0.7244, 0.1325) for the inclined wall and V-groove, respectively. We observe a maximum reduction in error of $5\times$ and an increase in temporal resolution of up to $30\times$.

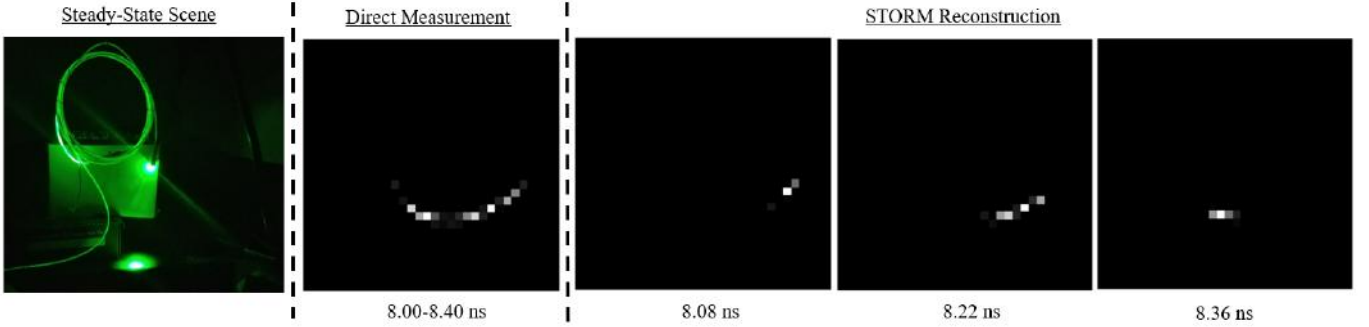


Fig. 7: **Transient reconstruction using STORM allows finer tracking of light in flight:** Single frame captures from direct measurement and STORM reconstruction of light-in-flight imaging. Image on the left shows the steady-state image capture of laser pulses (at 77.8 MHz repetition rate) flowing through a fiber optic looped five times in a ring of diameter 7 inches. The image next to the steady-state image is the direct measurement of the system at the time interval between 8.00 ns and 8.40 ns. The distance travelled by the pulse during this frame is about 5 inches. The three frames to the right show the result of STORM reconstruction during time points in the same time interval. The smallest localization of the pulse is about 1 inch, resulting in temporal resolution improvement of $5\times$.

where x is chosen to range between distances within the field of view of the SPAD with $x = 0$ located on the camera axis. For the V-groove scene, $-|x|$ is used since it has two incline planes moving closer to the detector. Since the initial depth of the scene (d_o) and angle the planes made with the camera axis (ϕ) were computed by hand, the ground truths computed here may have some error.

B. Light-in-Flight imaging

O'Toole et al. [35] demonstrated light propagation through an optical fiber. We performed transient imaging experiments with and without STORM, using a similar scene (Figure 7), created using a coiled fiber that is connected to the laser

and looped five times. The diameter of the resulting circle is approximately 7 inches. The fiber is intentionally not well coupled, and therefore light leaks from the cladding as the laser pulse propagates through the fiber. As the leaked photons reach the SPAD array and are time stamped, we can track the light traveling through the fiber. We perform an extra post-processing peak detection step for both direct and STORM reconstructed transients, to better visualize the pulse propagating through the loop. Figure 7 shows reconstructed frames for the time interval between 8.00 ns and 8.40 ns. In both direct measurements and STORM, a pulse can be seen moving through the fiber loop. However, STORM allows us to better visualize this propagation process, by increasing the resolution

of the captured transient video. The length corresponding to the pulse visible in the direct measurement frame is approximately 5 inches, whereas the analogous length in the STORM frame corresponding to time 8.36 ns is approximately 1 inch. This indicates an improvement in temporal resolution of $5\times$. The full video of the light traveling through the fiber loop is available in the supplement.

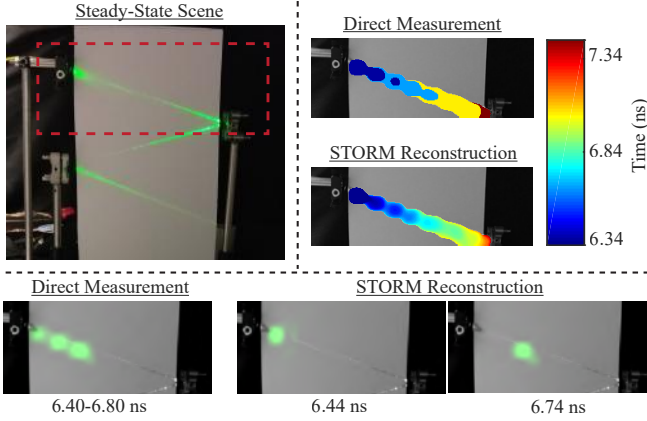


Fig. 8: Visualization of photon arrival times demonstrates improvement in temporal resolution of STORM: Scene in diagram to the left is of a laser bouncing off two mirrors. A white cardboard plane is oriented such that the laser glances off of it and the propagation of light can be imaged. The photon pulse is tracked within the red highlighted portion. Images in the top-right visualize the time of photon arrival along the path of propagation. This is demonstrated by the bottom 3 frames for both direct and STORM. There are 4 distinct time bins for direct measurement and 31 distinct time bins for STORM reconstruction resulting in an improvement by approximately $8\times$.

We perform another transient imaging experiment using a scene first used for light-in-flight visualization by Garipey et al. [13], shown in Figure 8. In [13], light-in-flight is imaged by the SPAD detector as the light scatters off air molecules. To increase the number of photons reaching the detector, we placed a white cardboard plane in the light path as shown in Figure 8(a). The still frames captured using the SPAD are superimposed on a 256×256 image of the scene. In the figure, different colors correspond to the photon arrival time along the path of propagation. The 3 frames in the bottom of Figure 8 display the photon pulse traveling. Using direct measurements, the photon path can only be discretized into 4 discrete time bins. STORM reconstruction improves the temporal resolution by approximately $8\times$ (31 time bins).

VII. CONCLUSIONS AND DISCUSSIONS

We introduced STORM, a transient super-resolution technique that uses multiple time-shifted measurements to overcome the limits on temporal resolution imposed by TDC quantization. Through proof-of-concept experiments using a real 2D SPAD array (with specifications representative of current commercially-available sensors), we demonstrated improvements in temporal resolution of at least $8\times$, and im-

provements in depth reconstruction error of at least $5\times$. This improved performance can immediately benefit a diverse set of applications, including ballistic imaging through scattering, non-line-of-sight imaging, and depth sensing. Additionally, the range of potential applications further increases, if we consider that STORM is applicable to other transient imaging technologies with jitter-to-bin-width characteristics similar to SPAD arrays, including streak cameras and ICCDs.

Moving forward, careful estimation of the system jitter $h(t)$ can lead to a more accurate estimate of the desired transient, $x(t)$. Instead of directly calibrating the jitter function, this could potentially be achieved by solving a *blind* deconvolution problem to simultaneously recover both $h(t)$ and $x(t)$. Finally, it will be interesting to explore combinations of STORM with orthogonal temporal super-resolution techniques that improve resolution through accurate modeling of noise characteristics in both the low-flux and high-flux regimes, or by exploiting spatial structure [19], [46].

VIII. ACKNOWLEDGMENTS

This work is supported by DARPA REVEAL (HR0011-16-C-0028) and NSF Expeditions (CCF-1730574, CCF-1730147) grants. Ankit Raghuram is also supported by a training fellowship from the Gulf Coast Consortia, on the IGERT: Neuroengineering from Cells to Systems, NSF (1250104). Ashok Veeraraghavan is also supported by NSF CAREER (IIS-1652633).

REFERENCES

- [1] N. Abaskharoun, M. Hafed, and G. W. Roberts. Strategies for on-chip sub-nanosecond signal capture and timing measurements. In *ISCAS 2001. The 2001 IEEE International Symposium on Circuits and Systems (Cat. No. 01CH37196)*, volume 4, pages 174–177. IEEE, 2001.
- [2] A. Agrawal and R. Raskar. Resolving objects at higher resolution from a single motion-blurred image. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007.
- [3] B. Bailey, D. L. Farkas, D. L. Taylor, and F. Lanni. Enhancement of axial resolution in fluorescence microscopy by standing-wave excitation. *Nature*, 366(6450):44, 1993.
- [4] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (9):1167–1183, 2002.
- [5] W. Becker. *Advanced time-correlated single photon counting techniques*, volume 81. Springer Science & Business Media, 2005.
- [6] M. Ben-Ezra, A. Zomet, and S. K. Nayar. Jitter camera: High resolution video from a low resolution detector. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–II. IEEE, 2004.
- [7] M. Ben-Ezra, A. Zomet, and S. K. Nayar. Video super-resolution using controlled subpixel detector shifts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):977–987, 2005.
- [8] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacio, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, 2006.
- [9] L. Cester, A. Lyons, M. Braidotti, and D. Faccio. Time-of-flight imaging at 10 ps resolution with an iccd camera. *Sensors*, 19(1):180, 2019.
- [10] S. Cova, A. Lacaita, M. Ghioni, G. Ripamonti, and T. Louis. 20-ps timing resolution with single-photon avalanche diodes. *Review of scientific instruments*, 60(6):1104–1110, 1989.
- [11] M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (slam) problem. *IEEE Transactions on robotics and automation*, 17(3):229–241, 2001.
- [12] Franco Zappa. 32×32 SPAD+TDC Camera datasheet. http://www.everyphotoncounts.com/files/Datasheet_32x32_SPAD+TDC_camera.pdf. Accessed: 03-27-2019.

- [13] G. Gariepy, N. Krstajić, R. Henderson, C. Li, R. R. Thomson, G. S. Buller, B. Heshmat, R. Raskar, J. Leach, and D. Faccio. Single-photon sensitive light-in-flight imaging. *Nature communications*, 6:6021, 2015.
- [14] I. Gkioulekas, A. Levin, F. Durand, and T. Zickler. Micron-scale light transport decomposition using interferometry. *ACM Transactions on Graphics (ToG)*, 34(4):37, 2015.
- [15] I. Gkioulekas, S. Zhao, K. Bala, T. Zickler, and A. Levin. Inverse volume rendering with material dictionaries. *ACM Transactions on Graphics (TOG)*, 32(6):162, 2013.
- [16] T. Gruber, F. Julca-Aguilar, M. Bijelic, W. Ritter, K. Dietmayer, and F. Heide. Gated2depth: Real-time dense lidar from gated images. *arXiv preprint arXiv:1902.04997*, 2019.
- [17] M. G. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of microscopy*, 198(2):82–87, 2000.
- [18] M. G. Gustafsson. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences*, 102(37):13081–13086, 2005.
- [19] F. Heide, S. Diamond, D. B. Lindell, and G. Wetzstein. Sub-picosecond photon-efficient 3d imaging using single-photon sensors. *arXiv preprint arXiv:1806.03381*, 2018.
- [20] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics (ToG)*, 32(4):45, 2013.
- [21] S. W. Hell and J. Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters*, 19(11):780–782, 1994.
- [22] Q. Hernandez, D. Gutierrez, and A. Jarabo. A computational model of a single-photon avalanche diode sensor for transient imaging. *arXiv preprint arXiv:1703.02635*, 2017.
- [23] J. Holloway, Y. Wu, M. K. Sharma, O. Cossairt, and A. Veeraraghavan. Savi: Synthetic apertures for long-range, subdiffraction-limited visible imaging using fourier ptychography. *Science advances*, 3(4):e1602564, 2017.
- [24] M. Janda and Z. Machala. Imaging of transient spark in atmospheric air by fast iccd camera. *IEEE Transactions on Plasma Science*, 39(11):2246–2247, 2011.
- [25] A. Jarabo, B. Masia, J. Marco, and D. Gutierrez. Recent advances in transient imaging: A computer graphics and vision perspective. *Visual Informatics*, 1(1):65–79, 2017.
- [26] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Transactions on Graphics (ToG)*, 32(6):167, 2013.
- [27] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar. Looking around the corner using ultrafast transient imaging. *International journal of computer vision*, 95(1):13–28, 2011.
- [28] P. Kramer, O. Hadar, J. Benois-Pineau, and J.-P. Domenger. Use of motion information in super-resolution mosaicing. In *Image Processing, 2006 IEEE International Conference on*, pages 357–360. IEEE, 2006.
- [29] C. Li, W. Yin, and Y. Zhang. Tval3: Tv minimization by augmented lagrangian and alternating direction algorithm 2009.
- [30] Z. Lin and H.-Y. Shum. Fundamental limits of reconstruction-based superresolution algorithms under local translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):83–97, 2004.
- [31] D. B. Lindell, M. O’Toole, and G. Wetzstein. Single-photon 3d imaging with deep sensor fusion. *ACM Transactions on Graphics (TOG)*, 37(4):113, 2018.
- [32] Micro Photon Devices. Psd-mod: Picosecond delay. <https://www.picoquant.com/products/category/accessories/psd-mod-picosecond-delayer>. Accessed: 2018-12-17.
- [33] K. Nasrollahi and T. B. Moeslund. Super-resolution: a comprehensive survey. *Machine vision and applications*, 25(6):1423–1468, 2014.
- [34] C. Niclass and E. Charbon. A single photon detector array with 64/spl times/64 resolution and millimetric depth accuracy for 3d imaging. In *Solid-State Circuits Conference, 2005. Digest of Technical Papers. ISSCC. 2005 IEEE International*, pages 364–604. IEEE, 2005.
- [35] M. O’Toole, F. Heide, D. B. Lindell, K. Zang, S. Diamond, and G. Wetzstein. Reconstructing transient images from single-photon sensors. In *CVPR*, volume 1, 2017.
- [36] M. O’Toole, D. B. Lindell, and G. Wetzstein. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature*, 555(7696):338, 2018.
- [37] A. Pediredla, A. Veeraraghavan, and I. Gkioulekas. Elliptic path sampling for time-gated rendering. *ACM Transactions on Graphics (TOG)*, 2019.
- [38] A. K. Pediredla, M. Buttafava, A. Tosi, O. Cossairt, and A. Veeraraghavan. Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner. In *2017 IEEE International Conference on Computational Photography (ICCP)*, pages 1–12. IEEE, 2017.
- [39] A. K. Pediredla, A. C. Sankaranarayanan, M. Buttafava, A. Tosi, and A. Veeraraghavan. Signal processing based pile-up compensation for gated single-photon avalanche diodes. *arXiv preprint arXiv:1806.07437*, 2018.
- [40] Pediredla, Adithya and others. Mitsuba time-of-flight renderer. <https://github.com/cmu-ci-lab/MitsubaToFRenderer>. Accessed: 12-15-2018.
- [41] S. Peleg, D. Keren, and L. Schweitzer. Improving image resolution using subpixel motion. *Pattern recognition letters*, 5(3):223–226, 1987.
- [42] C. Peters, J. Klein, M. B. Hullin, and R. Klein. Solving trigonometric moment problems for fast transient imaging. *ACM Transactions on Graphics (TOG)*, 34(6):220, 2015.
- [43] PicoQuant. Hydrharp 400. <https://www.picoquant.com/products/category/tcspc-and-time-tagging-modules/hydrharp-400-multichannel-picosecond-event-timer-tcspc-module>. Accessed: 2018-12-17.
- [44] PicoQuant. Picoharp 300. <https://www.picoquant.com/products/category/tcspc-and-time-tagging-modules/picoharp-300-stand-alone-tcspc-module-with-usb-interface>. Accessed: 2018-12-17.
- [45] J. Rapp, R. M. Dawson, and V. K. Goyal. Improving lidar depth resolution with dither. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 1553–1557. IEEE, 2018.
- [46] J. Rapp, Y. Ma, R. Dawson, and V. K. Goyal. Dead time compensation for high-flux ranging. *arXiv preprint arXiv:1810.11145*, 2018.
- [47] M. J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature methods*, 3(10):793, 2006.
- [48] G. Satat, B. Heshmat, D. Raviv, and R. Raskar. All photons imaging through volumetric scattering. *Scientific reports*, 6:33946, 2016.
- [49] O. Shahar, A. Faktor, and M. Irani. Space-time super-resolution from a single video. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3353–3360. IEEE, 2011.
- [50] E. Shechtman, Y. Caspi, and M. Irani. Space-time super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):531–545, 2005.
- [51] D. Shin. *Computational imaging with small numbers of photons*. PhD thesis, Massachusetts Institute of Technology, 2016.
- [52] R. Tadano, A. K. Pediredla, K. Mitra, and A. Veeraraghavan. Spatial phase-sweep: Increasing temporal resolution of transient imaging using a light source array. In *Image Processing (ICIP), 2016 IEEE International Conference on*, pages 1564–1568. IEEE, 2016.
- [53] L. Tian, X. Li, K. Ramchandran, and L. Waller. Multiplexed coded illumination for fourier ptychography with an led array microscope. *Biomedical optics express*, 5(7):2376–2389, 2014.
- [54] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature communications*, 3:745, 2012.
- [55] A. Velten, D. Wu, A. Jarabo, B. Masia, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, D. Gutierrez, and R. Raskar. Femto-photography: capturing and visualizing the propagation of light. *ACM Transactions on Graphics (ToG)*, 32(4):44, 2013.
- [56] A. Velten, D. Wu, A. Jarabo, B. Masia, C. Barsi, E. Lawson, C. Joshi, D. Gutierrez, M. G. Bawendi, and R. Raskar. Relativistic ultrafast rendering using time-of-flight imaging. In *ACM SIGGRAPH 2012 Talks*, page 41. ACM, 2012.
- [57] F. Villa, R. Lussana, D. Bronzi, S. Tisa, A. Tosi, F. Zappa, A. Dalla Mora, D. Contini, D. Durini, S. Weyers, et al. Cmos imager with 1024 spads and tdes for single-photon timing and 3-d time-of-flight. *IEEE journal of selected topics in quantum electronics*, 20(6):364–373, 2014.
- [58] K. Yu, N. Park, D. Lee, and O. Solgaard. Superresolution digital image enhancement by subpixel image translation with a scanning micromirror. *IEEE Journal of Selected Topics in Quantum Electronics*, 13(2):304–311, 2007.
- [59] L. Zelnik-Manor and M. Irani. Multi-frame estimation of planar motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1105–1116, 2000.
- [60] G. Zheng, R. Horstmeyer, and C. Yang. Wide-field, high-resolution fourier ptychographic microscopy. *Nature photonics*, 7(9):739, 2013.