Single-Photon Structured Light

Varun Sundar† Sizhuo Ma† Aswin C. Sankaranarayanan† Mohit Gupta†
† University of Wisconsin-Madison ‡ Carnegie Mellon University

Abstract

We present a novel structured light technique that uses Single Photon Avalanche Diode (SPAD) arrays to enable 3D scanning at high-frame rates and low-light levels. This technique, called “Single-Photon Structured Light”, works by sensing binary images that indicates the presence or absence of photon arrivals during each exposure; the SPAD array is used in conjunction with a high-speed binary projector, with both devices operated at speeds as high as 20 kHz. The binary images that we acquire are heavily influenced by photon noise and are easily corrupted by ambient sources of light. To address this, we develop novel temporal sequences using error correction codes that are designed to be robust to short-range effects like projector and camera defocus as well as resolution mismatch between the two devices. Our lab prototype is capable of 3D imaging in challenging scenarios involving objects with extremely low albedo or undergoing fast motion, as well as scenes under strong ambient illumination.

1. Introduction

Structured light (SL) 3D imaging systems have inherent tradeoffs that balance the precision of the 3D scan against its acquisition time. For instance, temporally-multiplexed SL techniques [21, 22, 25] achieve high depth resolution by projecting multiple patterns, thereby precluding high-speed capture. At the other extreme, SL based on spatially-modulated patterns [27, 62] can facilitate single-shot scans, but require assumptions of spatial-smoothness that invariably result in loss of detail.

The tradeoffs inherent to SL systems are exacerbated when operating in challenging regimes with low signal-to-noise ratios (SNR) arising from either low-albedo objects, dynamic scenes, or strong ambient illumination. In these scenarios, using longer temporal codes can offer robustness and precision, but at the cost of lowered acquisition speed. One way to mitigate the loss in time resolution is to use high-speed cameras and projectors. However, this approach is limited by large bandwidth requirements and, more fundamentally, the presence of read noise. Each image has a constant amount of read noise, immaterial of the exposure time; this can dominate the received signal as the exposure times and, consequently, the image intensities are reduced.

This paper envisions a class of Single-Photon Structured Light systems that are based on single-photon detectors, such as Single Photon Avalanche Diodes (SPADs). SPADs can be operated at very high speeds when detecting photons and not their time-of-arrivals. In this ‘photon detection’ mode, the measurements are binary-valued—indicating whether or not a photon arrival occurred during a given acquisition time. For instance, a recently developed SPAD array [60] can capture \(\sim 10^5\) binary frames at 1/8-th megapixel resolution. Our key observation is that the binary measurements, normally considered a limitation due to limited information, are sufficient for a large family of SL coding schemes [53] that are binary as well. Since SPADs count photon arrivals, they are not corrupted by read noise. Finally, the use of SPADs for SL finds a natural coupling in high-speed projectors that use digital micromirror devices (DMDs) for displaying binary patterns.

Coding and decoding for Single-Photon SL. Due to the probabilistic nature of photon arrivals, the binary-valued measurements captured by SPADs are prone to strong photon noise. For instance, in the presence of strong ambient light, the SPAD could detect a photon even when the corresponding projector pixel is dark. Traditional SL coding schemes are designed for regimes where the image measurements are not binary-valued, and hence are not suitable for Single-Photon SL. We formulate novel SL encoding strategies using error-correction codes that enable robust decoding for Single-Photon SL even under large photon noise.

Beyond achieving robustness to photon noise, SL coding schemes must account for various imaging phenomena such as projector and camera defocus. Naïve error-correcting codes do not consider these practical effects, and thus cannot be used in a real-world SL system. We design a new class of hierarchical codes using error correction and binary phase shifting that guarantee a minimum stripe width, which provides robustness to such non-idealities. Finally, we design a high-throughput decoding scheme for the proposed codes to enable real-time decoding of the measurements. Our implementation can decode a disparity map for
High Speed + High Precision + Low Albedo

Scene

Photon noise limited

DMD projector (Binary patterns) 20,000 Hz

SPAD array (Binary captures) 20,000 Hz

High Speed + High Precision + Low Albedo

Rapid hand motion

Depth (in mm)

520
565
605

Car Tire

250 FPS

80 FPS

40 FPS

Figure 1. Single-Photon Structured Light. (left) Our proposed system comprises of a SPAD array and a DMD-based projector used to project and acquire binary patterns at extremely high frame rates. We devise coding schemes that can obtain depth-maps from these photon noise limited captures. (right) We demonstrate several practical capabilities of Single-Photon SL, including: high-speed scanning of a rapidly moving hand at 250 FPS, sub-millimeter precise depth-imaging at a range 50 cm range and at 80 FPS, and reconstructing the tread pattern of a tire, a low-albedo object, at 40 FPS. We include RGB frames captured at 30 FPS to depict the high speeds involved.

a 512 × 256 array in 100 ms on a CPU and 3 ms on a GPU.

Implications. Single-photon SL has the potential to enable extreme 3D imaging capabilities, including high-speed 3D scanning and robust 3D imaging in low-SNR conditions while respecting low-power and latency budgets. Figure 1 demonstrates several unique practical capabilities of our prototype Single-Photon SL system, including scanning scenes with low albedo (a tyre) and at high frame rate (fast hand movements) with little loss in the spatial resolution.

Limitations. Single-Photon SL inherits limitations endemic to many SL systems. While we mitigate short-range effects such as defocus, long-range effects such as inter-reflections remain to be addressed, possibly by incorporating existing work addressing global illumination. Current SPAD technology is still nascent compared to its CMOS counterparts; the low-resolution of SPAD arrays and their poor fill factors constrains the reconstruction quality of our approach. Fortunately, the capabilities of single-photon sensors continue to improve with higher resolution arrays featuring increased fill factors [47,48] on the horizon.

2. Related Work

Structured light 3D imaging. Active triangulation techniques have a rich history with early techniques including stripe scanning [3,38,57], shadow scanning [7,8], binary patterns [53] and sinusoid patterns based phase-shifting [58]. Many methods achieve fast single-shot acquisition by projecting statistical patterns [34,55,63] or via Fourier Transform Profilometry (FTP) [4,32,33]. Such techniques require spatial-smoothness assumptions and have low accuracy for strongly textured surfaces.

Fast binary projectors in SL. Several SL systems achieve high-speed 3D scans [27,35,36,59,62,64] using the projection capabilities of DMDs. However, all of these techniques use sensors based on traditional photodiodes which, unlike SPADs, are fundamentally limited by read noise.

Event-based 3D imaging. Event-based cameras are bio-inspired devices [17] that are triggered asynchronously by intensity changes (or “events”) typically from a pulsed laser [9,43,45] or by a DMD projecting multiple patterns [26,37,42]. In contrast to SPADs, event-cameras have 1-2 orders of magnitude lower event density (∼10⁶ events/s [39]). Since each event recovers at most a single 3D point, the limited event density lowers the density and quality of the reconstruction especially in presence of scene-wide motion. While event-based cameras can achieve high dynamic range, their low-light sensitivity remains poor, precluding reliable 3D imaging in low albedo and low SNR scenarios.

Single-Photon imaging. Only recently have the capabilities of SPADs, operating without any temporal synchronization, been explored, with applications in high-dynamic range imaging [28,29] and burst photography [41]. Our method operates the SPAD array similar to Ma et al. [41], using a sequence of binary frames. Although we focus on SPADs due to their superior frame-rate, the proposed techniques are applicable to other single-photon imaging technologies such as Jots [14,15], which feature high-resolution arrays with smaller pixels and increased photon-efficiency [40], albeit at lower frame-rates and higher read noise.

3. Image Formation in Single-Photon SL

Consider a SPAD pixel array observing a scene. The number of photons \( N \) arriving at a pixel \( x \) during an expo-
Figure 2. **Single-Photon SL features an asymmetric noise model.** Bit-flip probabilities as determined by Eqs. (3) and (4) are evaluated across a grid of \((\Phi_a, \Phi_p)\) flux values. The plot parameters are \(t_{\text{exp}} = 10^{-4}\) s and dark current rate \(r_q = 0\).

Sure time \(t_{\text{exp}}\) is modeled as a Poisson random variable:

\[
\Pr \{ N = k \} = \frac{\Phi(x) t_{\text{exp}}^k e^{-\Phi(x) t_{\text{exp}}}}{k!},
\]

where \(\Phi(x)\) is the flux. During each exposure, a pixel detects at most one photon, returning a binary value \(B(x)\) such that \(B(x) = 1\) if the pixel detects one or more photons. Hence, \(B(x)\) is a Bernoulli random variable with

\[
\Pr \{ B(x) = 0 \} = e^{-(\Phi(x) + r_q) t_{\text{exp}}},
\]

where \(r_q\) is the dark current rate—the rate of spurious counts unrelated to incident photons.

In a typical SL scan, the scene is illuminated with a sequence of 2D binary patterns from a projector. The SPAD captures a binary frame for each pattern. Each SPAD pixel receives a binary code over time, from which we estimate the projector column observed at the pixel—an operation that is critical for the success of any SL technique.

We now derive the probability that a projected temporal sequence will be decoded incorrectly. For a given binary pattern, consider a SPAD pixel \(x\) that observes a scene point illuminated by an ON pixel. Suppose the incident photoelectron arrival rate at \(x\) due to projector and ambient illumination are \(\Phi_p(x)\) and \(\Phi_a\), respectively. Then, the probability of a bit-flip error, i.e., the probability of the SPAD pixel not detecting a photon is given as

\[
P_{\text{flip, bright}} = \Pr \{ B(x) = 0 \mid \Phi(x) = \Phi_a + \Phi_p(x) \}. \tag{3}
\]

Similarly, the probability of a bit-flip error when the projector pixel is OFF, i.e., the probability of detecting a photon in spite of not illuminating the corresponding projector pixel is

\[
P_{\text{flip, dark}} = \Pr \{ B(x) = 1 \mid \Phi(x) = \Phi_a \}. \tag{4}
\]

The bit-flip probabilities for “bright” and “dark” pixels, visualized in Fig. 2 for varying \(\Phi_a\) and \(\Phi_p\), are not equal due to the asymmetric role played by the ambient photons.

We can now compute the probability of incorrectly decoding an \(L\)-length Gray code. Since it is equally likely to observe any \(L\)-bit binary code, the average probability of erroneous decoding over all codewords is:

\[
\Pr \{ \text{error} \} = 1 - \left(1 - \frac{P_{\text{flip, bright}} + P_{\text{flip, dark}}}{2}\right)^L. \tag{5}
\]

A detailed derivation is provided in Suppl. Sec. 1.2.

**Typical decoding error probabilities.** Figure 3 shows the decoding error probability for a 10-bit Gray code across ambient light levels. Increasing ambient light levels drastically increases \(P_{\text{flip, dark}}\), resulting in a near-certain decoding failure. In the next section, we propose coding strategies for Single-Photon SL that enable accurate decoding even in highly challenging conditions.

### 4. Coding for Single-Photon SL

We now describe temporal coding schemes for Single-Photon SL with the goal of achieving robustness to random bit-flips using error correction mechanisms, and incorporating practical considerations in code design. The overall coding and decoding pipeline is illustrated in Fig. 4.

One simple strategy to improve the reliability of any scheme is to repeat the projected patterns and perform a majority vote. This seems to be a viable option since the high-speed projection and capture of DMDs and SPADs, respectively, affords high temporal redundancy. For example, given a code sequence of length \(L = 10\), we could simply repeat the patterns 25 times (called the redundancy factor), and still maintain a high overall frame rate for 3D capture. Such a majority vote will improve the decoding performance, provided the probability of bit flips is less than 0.5. However, in extreme conditions (say, low SNR) a large
A BCH\((n, k, d) : \{0, 1\}^k \to \{0, 1\}^n\) encoder takes input messages of length \(k\) and produces output codewords of length \(n\) that are at least \(d\)-bits apart. Hence, such a coding scheme provides error correction capabilities up to \(\lceil \frac{d-1}{2} \rceil\) bit flips, in the worst case or adversarial sense. This worst-case error-correcting capacity of BCH codes is significantly higher than of the repetition code. For instance, BCH\((63, 10, 27)\)—which uses 63-length codewords to encode messages of length 10—can correct up to 13 worst-case errors, while the corresponding capability for Repetition\((60, 10, 6)\), where the message pattern is repeated 6 times, can correct only 2 errors. Going further, in our problem setting, the main source of bit flips is photon noise, which is stochastic and non-adversarial, and thus we can expect error correction beyond the worst-case regime.

**Designing BCH-encoded patterns.** Consider a projector with \(C\) columns. We aim to design projector patterns so that each column is assigned a unique binary code with in-built BCH error correction. To produce the projector patterns, we start with a base binary coding scheme that uniquely represents each projector column, for example, with Gray codes\([30]\). Given a set of message codes \(\{m_i\} \subseteq \{0, 1\}^L\), where \(L = \lceil \log_2 C \rceil\), we choose a BCH encoder \(E_{\text{BCH}(n, k, d)}\) that is capable of encoding at least \(C\) messages (i.e., \(k \geq L\)) and output column-wise codes \(\{e_i \in E_{\text{BCH}(n, k, d)}(m_i)\}\).

As an example, Fig. 5 illustrates the BCH\((31, 11, 11)\) encoding of 10-bit Gray code messages. Since \(L < k\) here, we use shortening, i.e., we prepend the message by \((k - L)\) zeros, but do not transmit them, thereby reducing the projected code length from \(n\) to \(n - (k - L)\). We also use systematic encoding, i.e., the first \(L\)-bits of each code is the message itself—hence each sequence comprises of message patterns appended by parity patterns. The choice of \(n\), the length of the BCH code, simultaneously determines the robustness of the code as well as the loss in time resolution. Longer codes have better error-correction capabilities; but since we need to acquire a larger temporal sequence, this reduces our ability to handle fast(er) moving objects. With this in mind, in the rest of the paper, we present results at two operating points—\(n = \{63, 255\}\)—to cover two distinct scenarios.

**Evaluating BCH encoding for Single-Photon SL.** To un-
understand the benefits of BCH encoding, we use Monte-Carlo simulation of decoding error probability across a grid of \((\Phi_a, \Phi_p)\) values, and compare conventional Gray codes, repetition codes and BCH codes for 10-bit binary messages. The performance of these schemes is presented in Fig. 6. At most operating points, the decoding error probability of BCH codes is either close to zero or presents an order of magnitude improvement over repetition.

4.2. Code Design under Practical Considerations

Beyond achieving robustness to photon noise, SL coding schemes must account for various imaging non-idealities. A majority of the BCH encoded patterns, as seen in Fig. 5, comprise of high-spatial frequency patterns that do not perform well under projector/camera defocus and resolution mismatch between the devices. Therefore, in spite of achieving low errors in theory, BCH codes as described so far will simply be inadequate in a practical SL system.

A common approach to mitigate such short-range effects is to use long-run Gray codes, which are a subset of Gray codes that maximize the shortest stripe width [18, 19, 23] across all the projected patterns. However, applying BCH encoding on long-run Gray codes also results in a majority of patterns containing high-spatial frequencies (see Suppl. Fig. 4). Finding binary messages \(\{m_i\}\) that maximize the minimum stripe-width of BCH patterns \(\{c_i\}\) is an intractable combinatorial problem with an exorbitant solution space (1024! candidate solutions).

**Hybrid codes.** Our key idea is to design hierarchical codes where BCH encoding is performed only on the more significant bits (MSBs) of the base Gray code pattern. This ensures that all the BCH-encoded frames have large minimum stripe widths, making them robust to defocus effects. The remaining lower significant bits (LSBs) are resolved using circularly-shifted binary patterns, where we shift the pattern one-pixel-at-a-time to represent columns sequentially. We term this as “binary shifting”.

Specifically, given a \(L\)-bit message, we encode its \(L_{\text{BCH}}\) MSBs and the remaining \(L_{\text{shift}} = L - L_{\text{BCH}}\) LSBs in different ways. The MSBs are coded using BCH as described earlier; since the message codes corresponding to a specific MSB pattern remains unchanged for all values of the LSBs, the resulting BCH codes have a stripe width of at least \(2^{L_{\text{ain}}.}\)

The \(L_{\text{shift}}\) LSBs are coded by a temporal sequence of length \(2^{L_{\text{ain}}+1}\) featuring a burst of \(2^{L_{\text{ain}}}\) ones—whose starting position (or phase) encodes the message. Figure 8 illustrates the codewords arising for this hybrid construction, which are guaranteed to have an overall stripe width of at least \(2^{L_{\text{ain}}}\) pixels. In our implementation, for a \(L = 10\)-bit message, we set \(L_{\text{BCH}} = 7\) and \(L_{\text{shift}} = 3\), featuring a minimum stripe width of \(2^3 = 8\) pixels. Additionally, we utilize BCH encoders with \(n \in \{63, 255\}\).

We note that this coding scheme is similar in spirit to hybrid SL techniques [11, 20, 64, 65] where Gray codes provide global disambiguation and Phase Shifting [21] resolves LSBs, providing precise correspondences. However, binary shifting has a key difference compared to phase shifting in that intensity information cannot be inferred from a binary measurement. Consequently, unlike phase shifting, where a single measurement can determine the unwrapped phase, binary shifting requires projecting multiple patterns.

Binary shifted patterns are decoded using a matched filter approach, by autocorrelating the received sequence with an unshifted stripe sequence. In Suppl. Sec. 1.5, we show that binary shifted patterns offer significant robustness to random bit-flips by deriving the expected decoding error. Finally, to illustrate the hybrid codes’ overall error-correcting capability, we compare them to repeated long-run Gray codes of similar codelength. We characterize performance using root mean squared error (RMSE) in decoded correspondence as the error metric. As we observe in Fig. 8, hybrid codes outperform repeated Gray codes, and the performance gap increases at higher redundancy factors.

![Figure 5. BCH(31, 11, 11) encoding of a 10-bit Conventional Gray message. We use systematic encoding, where message patterns are appended by parity patterns, providing tolerance to random bit-flips caused by photon noise. The complete code lookup table describing these patterns is shown in Suppl. Fig. 3.](image)

![Figure 6. Monte-Carlo evaluation of BCH and repetition strategies. We empirically evaluate the probability of decoding error upon receiving a codeword randomly corrupted by bit-flips. The ambient flux \((\Phi_a)\) and projector flux \((\Phi_p)\) values at a pixel location determine the bit-flip probability. We use BCH encoders with \(n = \{63, 255\}\). Both repetition and BCH strategies improve the robustness of conventional Gray codes to photon noise. Additionally, BCH outperforms repetition at all \((\Phi_a, \Phi_p)\) with a pronounced difference at higher redundancies.](image)
RMSE and MDD being the Maximum Likelihood Decoder when $P$ is insufficient due to a potentially large number of bit-flips caused by photon noise. In this section, we discuss fast decoding techniques for BCH and, as an extension, the hybrid codes proposed in Sec. 4. The goal is to design techniques that can achieve real-time decoding, while also being able to handle a large number of individual bit-flip errors.

**Minimum distance decoding for Single-Photon SL.** One simple decoding approach is Minimum Distance Decoding (MDD), where the measured codewords are compared against every projected code word. MDD, while conceptually simple, can correct errors beyond the worst-case limit\(^2\). However, a brute-force implementation of MDD can often be unviable, owing to its exorbitant run-time and/or memory requirements. Fortunately, Single-Photon SL has certain favourable properties that lead to a fast, high-throughput MDD procedure. First, the space of messages (number of projector columns, $\sim 2^{10}$) is significantly smaller than space of codewords ($2^n$, $n \in \{63, 255\}$). Second, the number of queries for decoding, which is the number of pixels in the SPAD sensor, exceeds the number of messages.

These circumstances permit us to leverage the recent progress in similarity search \([5, 56]\), which has lead to efficient nearest-neighbour algorithms for batched queries. Based on empirical comparisons (presented in Suppl. Fig. 7), we find that FAISS \([31]\) offers the highest throughput, decoding a 1/8th MP array in 100 ms on CPU and 3 ms on GPU. Such methods also scale to larger arrays, requiring 12 ms and 30 ms for one and four megapixels respectively.

### 5. Experimental Results

We now describe a range of experiments to demonstrate the performance of Single-Photon SL. Our lab prototype was constructed with the SwissSPAD2 array \([60]\), which is a $512 \times 256$ SPAD array. The array has a pixel pitch of 16.38 $\mu$m, and can capture binary frames at speeds up to 100 kHz. In Suppl. Sec. 4, we provide additional details regarding the setup including the calibration procedure used.

#### 5.1. Single-Photon SL on Static Scenes

To characterize the performance of Single-Photon SL, we image static scenes of varying albedo and ambient light levels. We use these case studies to compare the performance of different error correction schemes and to show the effectiveness of the proposed hybrid codes. To obtain ground truth scans, we operate a DMD projector at a low frame-rate of 2 Hz, while running the SPAD at 10240 Hz, thereby obtaining 5120 SPAD frames per projected pattern. The average of 5120 frames has minimal photon noise, and is considered as a ground truth measurement.

**Performance of proposed codes.** Figure 9 compares our proposed Hybrid and BCH strategies to repeated Gray codes and repeated long-run Gray codes. We report overall RMSE and RMSE among inliers, thereby measuring both accuracy and consistency. As seen in Sec. 4.2, the proposed hybrid codes are considerably more consistent and accurate across the two redundancy factors used. Whereas, naive
Figure 9. **Strategy comparison for Single-Photon SL on porcelain bust.** *(left)* We obtain ground truth by averaging the burst of 5120 binary frames captured for each projected pattern. To illustrate the challenge of photon noise, we include reconstruction using Gray code without repetition that has severe artifacts. *(right)* Comparison between our proposed Hybrid strategy and other baseline methods across operating points $n = \{63, 255\}$. BCH codes have several high spatial-frequency frames, and are easily distorted by short-range effects. We report three metrics: RMSE in estimated depth, percentage of inliers (absolute depth error $< 5$ mm), and RMSE among these inliers.

BCH codes are heavily distorted due to defocus. Since each strategy has access only to a single binary frame per projected pattern, this emulates a 3D capture speed of 40 FPS in Hybrid ($n = 255$) and 130 FPS in Hybrid ($n = 63$).

**Low SNR regimes.** Figure 10 examines reconstruction quality across various ambient light sources, including indoor lighting and a bright work lamp. The reconstructions, shown for Hybrid ($n = 255$), are robust to ambient light, albeit with a drop in performance under the work lamp. These results can potentially be improved by judicious use of light redistribution schemes [24, 45, 49] and exposure control.

Next, we consider low-albedo scenes by imaging objects covered by highly absorptive materials, such as 3M Black Matte [1] and Acktar Velvet [2]; the latter absorbs up to 99.9% of incident light. As Fig. 12 shows, Single-Photon SL can recover the 3D geometry of these dark objects, even when visually imperceptible. As a practical example, Fig. 1 shows the reconstruction of tire treads scanned at 40 FPS.

**5.2. Dynamic Scenes with a High-Speed Projector**

For dynamic scenes, we used a projector based on the Texas Instruments DLP6500 DMD, capable of projecting binary images with a resolution of $1024 \times 768$ pixels at 20 kHz. For simplicity, we operate the SPAD at the same speed as the projector. The projector uses a broadband white LED (SugarCUBE Ultra White LED) as the illumination source.

Figure 13 shows high-speed 3D imaging for a sequence of fast hand movements. A commercial 3D scanner (Kinect-v2 camera [63]) operating at 30 FPS fails to recover the fingers of the rapidly moving hand, while Single-Photon SL continues to recover fine details. Finally, in Fig. 11, we reconstruct the deforming folds of a cloth as it is waved in front of the camera. For both sequences, we use Hybrid ($n = 63$) operated at 250 FPS. These demonstrate the ability of proposed Single-Photon SL techniques to recover detailed 3D geometry of high-speed deformable objects.
Figure 11. **Non-rigid deforming object captured by Single-Photon SL using Hybrid** \((n = 63)\) at 250 FPS. We include (a) a reference image captured by the SPAD camera using a long integration time, (b) a single binary frame and (c) the reconstructed meshes clearly showing the folds of the cloth. Capturing a non-rigid object is particularly challenging—unless we operate at high speeds, excessive motion blur is induced. **We include a high-speed depth video of this sequence in the supplementary material.**

![Reference image](image1.png)
![Binary Frame](image2.png)
![3D Reconstruction](image3.png)

(a) 3M Black Matte (absorbance 99%)
(b) Acktar Metal Velvet (absorbance 99.9%)

Figure 12. **3D reconstruction in low-albedo scenes** of (a) a white cylinder covered by 3M Black Matte and (b) an inverted V-groove covered by Acktar Metal Velvet; the materials used have extremely high absorbance of 99% and 99.9%, respectively. Both reconstructions are obtained using Hybrid \((n = 255)\) at 40 FPS.

In summary, our results on these challenging scenes—both static and dynamic—illustrates the practical capabilities of the Single-Photon SL modality and its ability to simultaneously achieve high speed, precision and robustness.

6. Conclusion and Discussion

This paper shows that many of the tradeoffs inherent to SL systems can be addressed via the use of SPAD sensors. Single-Photon SL, the system that we propose, exploits the single photon detection capabilities of SPAD sensors, along with its lack of read noise. The proposed ideas are capable of detecting objects with high absorbance, and scenes with dramatic high-speed motion. The enabling techniques underlying these are a set of error correction codes, that are designed to be resilient to aberrations commonly present in SL systems. As with many recent efforts in this space, Single-Photon SL provides yet another case study for the wider adoption of SPAD sensors in the imaging community.

**Optimal coding and decoding.** Despite their effectiveness, the proposed coding and decoding strategies are not provably optimal. Our MDD implementation is agnostic to the asymmetry of bit-flips, a defining feature of Single-Photon SL. While maximum likelihood decoders remain to be constructed for the general asymmetric case, optimal decoders have been derived for edge cases [13], e.g., when \(P_{\text{flip, bright}} >> P_{\text{flip, dark}}\). Further, leveraging an optimization framework [46] such as optical SGD [10] can lead to improved coding and decoding schemes for Single-Photon SL.

**Handling ambient illumination.** In strong ambient illumination, bit-flips arising from \(P_{\text{flip, dark}}\) are predominant. An important next step is to exploit such asymmetry of bit-flips to design Single-Photon SL codes that are optimized for extreme ambient illumination. Finally, another promising research direction is to explore Single-Photon SL with complementary modalities such as light concentration [24, 50] and epipolar structured light [51, 52] to gain further robustness to extreme ambient and global illumination.
References


[50] Matthew O’Toole, Supreeth Achar, Srinivasa G Narasimhan, and Kiriakos N. Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. ACM SIGGRAPH, (0), 2015. 8
[54] Ron M Roth. Introduction to coding theory. IET Communications, 47:18–19, 2006. 4


