

FAST VISION IN THE DARK: A CASE FOR SINGLE-PHOTON IMAGING IN PLANETARY NAVIGATION

David Rodríguez-Martínez* and C.J. Pérez del Pulgar

Space Robotics Lab, Dept. of Systems Engineering and Automation, University of Malaga,

**corresponding author: david.rm@uma.es*

ABSTRACT

Improving robotic navigation is critical for extending exploration range and enhancing operational efficiency. Vision-based navigation relying on traditional CCD or CMOS cameras faces major challenges when complex illumination conditions are paired with motion, limiting the range and accessibility of mobile planetary robots. In this study, we propose a novel approach to planetary navigation that leverages the unique imaging capabilities of Single-Photon Avalanche Diode (SPAD) cameras. We present the first comprehensive evaluation of single-photon imaging as an alternative passive sensing technology for robotic exploration missions targeting perceptually challenging locations, with a special emphasis on high-latitude lunar regions. We detail the operating principles and performance characteristics of SPAD cameras, assess their advantages and limitations in addressing key perception challenges of upcoming exploration missions to the Moon, and benchmark their performance under representative illumination conditions.

Key words: lunar exploration; vision-based navigation; advanced sensing; single-photon; low-light; HDR.

1. INTRODUCTION

Enhanced sensing and imaging systems are essential to the success of a number of high-priority planetary exploration missions, particularly those targeting environments where visual perception is severely impaired. Examples include prospecting at high-latitude lunar locations, characterized by low solar elevation angles and shorter diurnal cycles, or navigating the pitch-black interiors of Martian lava tubes. These environments present extreme illumination conditions—ranging from partial to complete darkness—paired with uneven, hazard-strewn landscapes. Navigating autonomously in such conditions demands high visual acuity, while also conserving limited onboard power and computational resources.

Vision-based navigation relying on Charge-Coupled Device (CCD) and Complementary Metal-Oxide-Semiconductor (CMOS) cameras faces significant challenges under complex and dynamic lighting conditions, which limit the operational range and accessibility

of planetary robots. For example, martian rovers can typically drive for only about 3 hours per sol when illumination conditions are optimal [1]. Similarly, NASA's *Ingenuity* helicopter was unable to fly when atmospheric conditions on Mars were best suited for it (dawn and dusk), since the insufficient illumination prevented its visual odometry from reliably detecting and matching ground features [2].

To address the lack of illumination, inconsistent brightness, and faster-than-usual traveling speeds [3, 4], high-latitude lunar exploration and prospecting missions must rely on either more robust computational image-processing algorithms (e.g., [5]), expand exteroceptive sensor suites to include a wider range of imaging modalities [6], employ improved and more reliable sensing technologies [7, 8], or, most likely, a combination thereof.

Among alternative sensing technologies, neuromorphic (event-based) cameras [7], thermal-infrared sensors [9], and active scanning Light Detection and Ranging (LiDAR) sensors [10] are presented as promising alternatives for enabling robust navigation in perceptually challenging environments, outperforming conventional optics in dimly lit, high dynamic range (HDR), and fast-moving scenarios. LiDAR, in particular, has already presented promising results for precise mapping of celestial body surfaces [11, 12], spacecraft navigation and landing [13], and autonomous in-orbit rendezvous and docking maneuvers [14]. But their application and proven performance in planetary exploration remains limited, largely due to the current, though rapidly maturing, technology readiness of solid-state, flash, and global shutter LiDARs [10]. Additional constraints include high power demands, mass and size constraints, limited range and restricted fields of view, and substantial data throughput requirements, which together demand careful engineering trade-offs before such systems can be reliably integrated into planetary missions. Notably, what characterizes all of the aforementioned alternative technologies is their distinct way of parsing reality, making it often necessary to pair them with complementary optical sensors to provide the complete contextual information most missions demand.

In this work, we introduce a novel imaging technology with a unique potential to enhance vision-based planetary

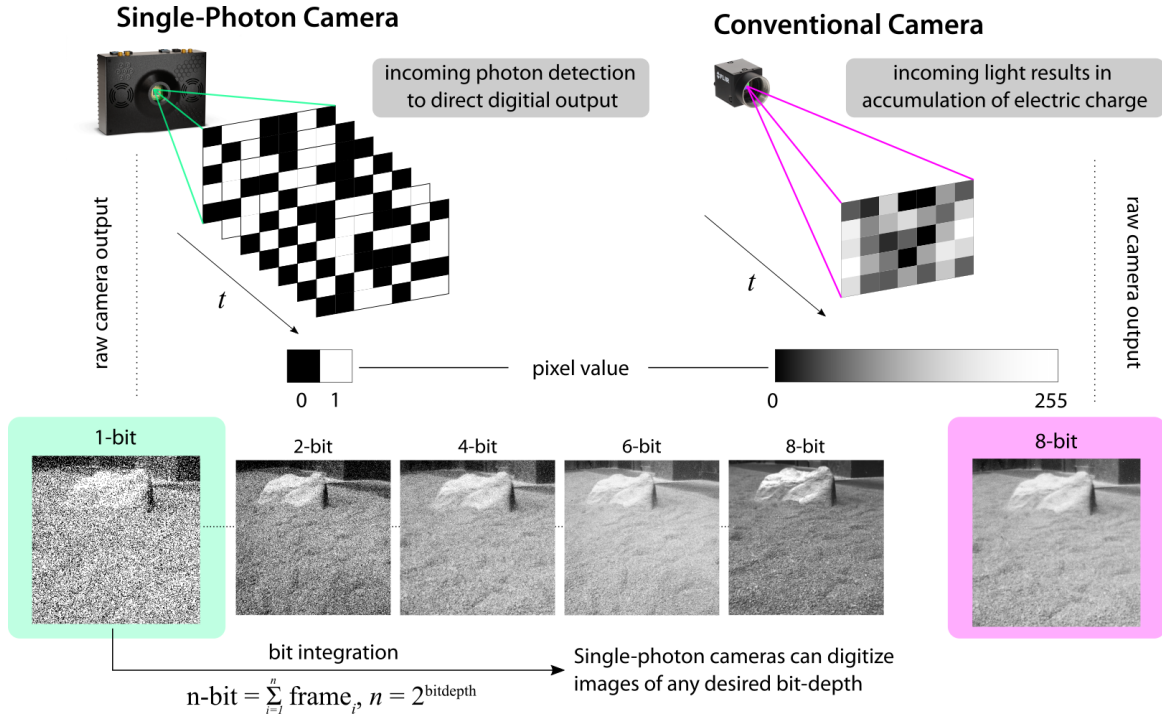


Figure 1: Unlike conventional cameras, which accumulate electric charge in each pixel over a given amount of time and subsequently convert it to a digital value (often of 8 bits per pixel), single-photon cameras provide a direct and synchronized digital output in the form of binary frames (1 bit per pixel). Each pixel value (0 or 1) indicates the absence or presence of an incident photon. This photon-level sensitivity enables the generation of images with arbitrary bit-depth by integrating 2^n number of binary frames, where n is the desired bit-depth of the final image. SPADs can also capture these binary frames at very high speeds, up to 100 kfps (10μ per frame), or even combined binary frames taken at different exposures to create high dynamic range images (>100 dB).

robotic systems due to its light sensitivity, speed, adaptability, and, despite its distinctive way of capturing visual information, its ability to operate and record visual data without the need for complementary sensors: **single-photon imaging**. Our main contributions include:

1. **Introduction of single-photon imaging for planetary navigation.** We focus on Single-Photon Avalanche Diode (SPAD) cameras, summarizing their operating principles, modeling methods, key characteristics, and assessing their advantages and limitations in addressing critical perception challenges of upcoming exploration missions.
2. **First evaluation of SPAD cameras as a passive imaging alternative for planetary navigation.** We compare their imaging and task-specific performance with a standard monochrome camera under lunar-analog conditions replicating the visual complexity and variable illumination of high-latitude regions.
3. **Identification of key challenges and future research directions.** We outline key obstacles to increasing SPAD technology readiness and enabling its integration into future missions, as well as potential benefits and unexplored research avenues.

2. SINGLE-PHOTON IMAGING

2.1. Operating Principle

Single-Photon Avalanche Diode (SPAD) cameras are a specific type of avalanche photo-diodes (APDs) capable of detecting the arrival of individual photons. They do so by operating in the so-called Geiger mode; i.e., at a high reverse bias voltage (above breakdown) where the strike of a single photon triggers an avalanche multiplication of charge carriers in the semiconductor, producing a distinct electrical pulse as an output [15]. This pulse is amplified directly within the SPAD pixel, providing a ready-to-use, inherently digital signal in the form of a synchronous binary stream where 0 means no photon detected and 1 means a photon was detected during that time window (see Figure 1). This capability, together with active quenching (rapid lowering of the bias voltage below breakdown to stop the otherwise self-sustaining avalanche process), enables SPADs to detect photon strikes with timing resolutions in the sub-10 picosecond range [16] and exceptionally low noise levels [17]. Note that, unlike other forms of photon-detection such as Quanta Image Sensors (QIS) based on extremely small specialized pixels called jots [18], SPADs are capable of detecting single-photon events but cannot resolve the exact number of photons arriving simultaneously within each detection time window [19].

The implementation of SPADs in CMOS technology [20] has facilitated their miniaturization and integration into large pixel arrays (>1Mpx, e.g., [21, 22]). This advancement has enabled their use as extremely fast (up to 100 kfps) and light-sensitive intensity imaging sensors, making SPADs suitable for a variety of applications, including robotics. Unlike conventional cameras, SPADs produce a noiseless, constant stream of binary frames. These frames can be accumulated and integrated to produce images of any desired bit depth (Figure 1). Even binary frames captured at different exposure times can be combined to produce high-speed HDR images.

These capabilities enable resolving visual features without the need for artificial lighting or auxiliary sensors, improving adaptability and robustness under extreme conditions (e.g., mx illuminance) while optimizing system design trade-offs. SPADs are already being used in a wide variety of applications, from 2D picosecond-resolution biomedical imaging [23] to active 3D time-of-flight (ToF) sensing and LiDAR [24]. Their rapid development has also drawn attention in the computer vision community ([25, 26, 27], to name a few). But their application and potential added-value have been seldom explored in the context of robotics and space [28, 29].

2.2. Imaging Model

The arrival of photons at a single SPAD pixel can be modeled as a random process. Assuming constant illuminance (and photon independence), the probability of a number of photons, k , striking a pixel within an exposure window follows a Poisson distribution [30]:

$$P(x = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad (1)$$

where λ is the expected number of photons, determined by the photon flux, ϕ (photons/second), the sensor quantum efficiency, η , and the exposure time, as in τ : $\lambda = E(x) = \phi\eta\tau$. If additional effects, such as dark counts (false detections in SPADs due to thermal noise triggering avalanches without the presence of incident photons), λ can also be defined as $\lambda = (\phi\eta + r_d)\tau$, where r_d indicates the dark count rate (DCR) in counts per second per pixel. In conventional sensors, this Poisson-distributed count directly determines the pixel intensity. For SPADs, however, the binary detection nature of every exposure means each pixel cannot count beyond 1 in a single frame, returning as a result a value of 1 if $k \geq 1$ or a value of 0 if $k = 0$. This detection is, therefore, governed by discrete Bernoulli statistics, where:

$$P(x = 0) = e^{-\lambda} = e^{-(\phi\eta\tau)} \quad (2)$$

$$P(x = 1) = 1 - P(x = 0) = 1 - e^{-(\phi\eta\tau)} \quad (3)$$

Over N binary frames (assuming no motion and perfect alignment), the maximum likelihood estimate (MLE) of $P(x = 1)$ per frame is simply defined by $\hat{p} = n/N$, where n is the number of frames with $k \geq 1$. One can,

therefore, estimate the MLE of the photon arrival rate per pixel, $\hat{\phi}$, as:

$$\hat{\phi} = \frac{1}{\eta\tau}(-\ln(1 - \hat{p})) = \frac{1}{\eta\tau}(-\ln(1 - \frac{n}{N})). \quad (4)$$

This statistical framework enables intensity reconstruction from sequences of binary frames.

3. METHODOLOGY

We grouped our evaluation of the performance of a single-photon camera against that of a standard monochrome camera into two categories. First, we assessed the **imaging performance** of both cameras under a range of illumination conditions—dawn/dusk, noon, and night—corresponding to illuminance levels from 100 lx down to 0.1 lx, at equivalent camera exposure settings. These conditions were specifically designed to simulate lighting conditions at the lunar poles, where Sun elevation angles range between 4° and 8° during daylight. Second, we assessed the **task-specific performance** of both cameras on mission-relevant tasks, including (A) scene segmentation, (B) lander detection, and (C) rover localization. For the first two, performance is evaluated over the results of the Segment Anything Model (SAM) [31], a transformer-based segmentation framework capable of performing zero-shot object and region segmentation across diverse image domains without task-specific training. We specifically implemented the pre-trained ViT-H model. For localization, we run our data on the well-known ORB-SLAM3 [32] pipeline, a feature-based, real-time visual-inertial SLAM system that supports monocular, stereo, and inertial input. These evaluations allowed us to compare not only raw imaging quality but also the practical impact of each camera on key operational tasks.

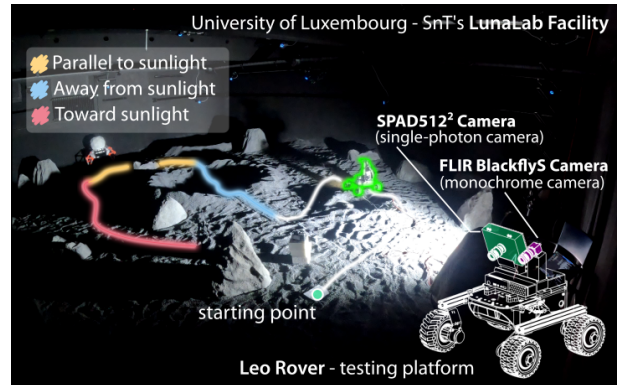


Figure 2: Experimental setup for our tests at the LunaLab.

3.1. Sensors

We used a SPAD512 single-photon camera from Pi Imaging Technology and a FLIR Blackfly-S from Teledyne Vision Solutions as our standard monochrome camera. Although the Blackfly-S is not a space-qualified camera, it is widely used in robotic vision applications, making it a

Table 1: Conventional vs single-photon camera specifications.

	FLIR Blackfly-S	SPAD 512
Sensor size	6.23×4.98 mm	9.5×9.5 mm
Resolution	720×540	512×512
Pixel Pitch	6.9 μm	16.38 μm
Bit Depth (raw)	8-bit	1-bit
EFL	14.24 mm f1.4	14.88 mm f1.4
FoV	21.6°×12.3°	35.5°×35.5°

EFL: Effective Focal Length; FoV: Field of View

relevant baseline for comparison. Optics for both cameras were chosen to account for their respective sensor sizes, ensuring similar equivalent focal lengths. The key specifications of each sensor are summarized in Table 1. Both cameras were mounted on a Leo Rover testing platform as shown in Figure 2. The rover also incorporated headlights to artificially illuminate the scene.

3.2. Data

The images used in our evaluation can all be found in the SPICE-HL3 dataset [33], recorded at the LunaLab of the University of Luxembourg-SnT. This facility is designed to replicate, with high photometric fidelity, lunar lighting conditions at various latitudes (Figure 2). The facility uses an Aputure LightStorm 600c Pro 600W spotlight, whose vertical and horizontal position, color temperature (2,300–10,000 K), and illuminance (69–51,100 lx) can be precisely adjusted to simulate different solar elevation angles and daylight conditions.

For the imaging performance evaluation, we primarily use images from the SPICE-HL3 static trajectory A. Data are divided into three main sections based on the heading angle with respect to the sun-simulator spotlight. These are referred to as *parallel to sunlight*, *away from sunlight*, and *toward sunlight* throughout the text. Visual odometry performance is evaluated on trajectories B, C, and D, each of which also corresponds to each of the previous phase angles, with the rover moving at 0.5 m/s. Because the native image resolutions of the SPAD and conventional camera differ, images are cropped to match FLIR’s aspect ratio of 4:3 for imaging performance, scene segmentation, and lander detection, and the SPAD’s 1:1 aspect ratio for visual odometry. To ensure comparability and compatibility with the implemented algorithms, most evaluations use SPAD images digitized at 8-bit. Although for the scene segmentation, we additionally look into the performance of the SPAD at a lower bit depth of 4 bits per pixel. No image enhancement or additional preprocessing is applied to either the SPAD or the FLIR data.

3.3. Metrics

Imaging performance was assessed using both qualitative and quantitative measures. Quantitative evaluation included contrast (RMS of pixel intensities), entropy (randomness of the intensity distribution, with higher values indicating greater detail), and sharpness (based on the variance of the Laplacian response). We also com-

puted BRISQUE and MS-SSIM scores. BRISQUE, or Blind/Referenceless Image Spatial Quality Evaluator, is a no-reference metric that models natural scene statistics to estimate perceived distortion (lower scores indicate better quality). MS-SSIM, or Multi-Scale Structural Similarity Index, is a full-reference structural similarity measure that operates across multiple resolutions (score range 0–1, with 1 indicating perfect agreement with a ground truth image). For MS-SSIM, images captured under regular laboratory illumination served as the reference.

For scene segmentation and lander identification, we evaluated the mean Intersection over Union (mIoU), inference time, and number of masks generated, with the latter two providing preliminary insights into computational demand variations across input images. In the visual odometry experiments, we measured the RMSE of the Absolute Trajectory Error (ATE) and the maximum drift (maximum ATE) for each trajectory.

4. RESULTS AND DISCUSSION

4.1. Imaging performance

The first step of our evaluation involved a direct visual assessment of the captured images to identify qualitative differences between the two camera types. Figure 3 presents example frames from each of the three distinct sections along trajectory A at different illumination conditions. All images evaluated were acquired with the rover headlights switched on. For the reference images, exposure times were chosen to yield well-balanced histogram distributions (i.e., mean intensity values near the midpoint of the dynamic range without clipping pixels on either edge). For the remaining images, exposure times were selected from those recorded under identical settings for both cameras, ensuring a fair basis for comparison.

The single-photon camera appears to consistently capture more detail of the scene under low-light conditions compared to the conventional camera. Under reference illumination, both camera types produce comparable images, with the conventional camera presenting a higher contrast and sharpness, likely due to its slightly higher resolution and efficiency at large photon counts. But as illumination decreases, the conventional camera shows a significant loss of detail, particularly at night, requiring a much longer exposure to resolve equal levels of detail in the image. In contrast, the single-photon camera maintains visibility of ground textures and obstacle features across all lighting conditions, capturing a better balance between lights and shadows, despite a clear lack of contrast.

Quantitative evaluation of image quality metrics across illumination conditions and exposure times highlights clear performance differences between the two camera types. The single-photon camera consistently outperforms the conventional camera in low-light scenarios (dawn/dusk and night) across most metrics, often achieving superior results at equal or substantially shorter exposure times. This advantage persists regardless of artificial

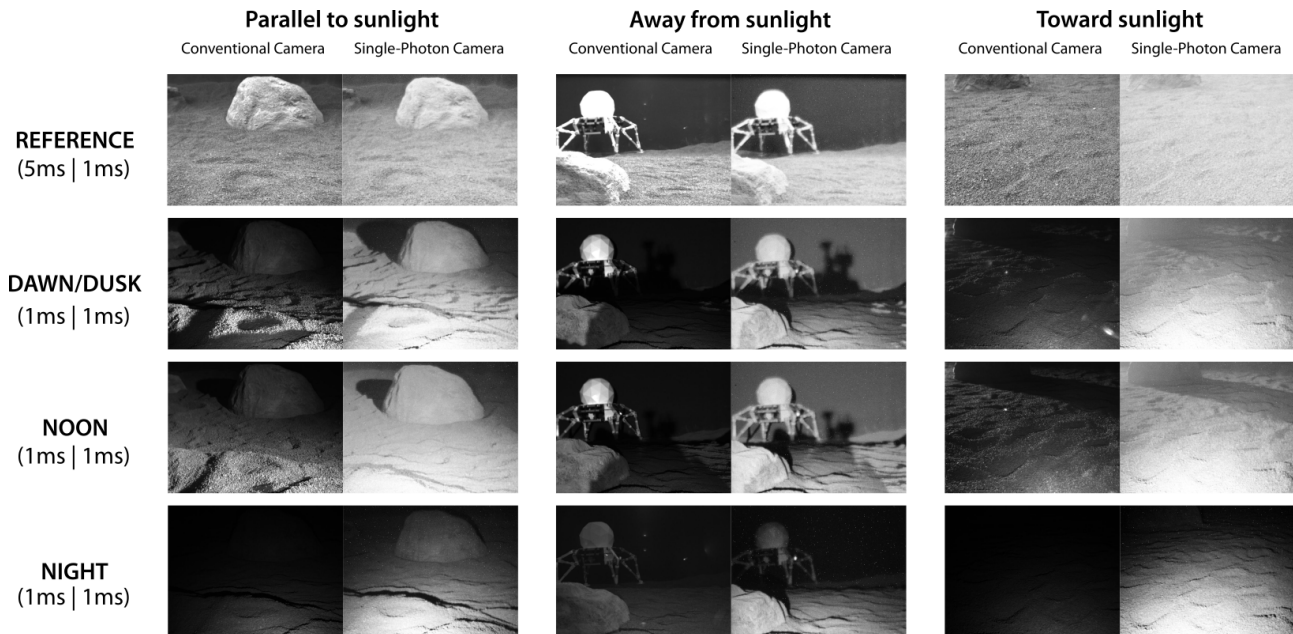


Figure 3: Selected frames taken by a conventional camera and a single-photon camera under different illumination conditions and at 8-bit equivalent exposure times. These frames are extracted from different sections of the SPICE-HL3 dataset trajectory A [33]. All images were captured with the rover headlights switched on.

illumination, underscoring the SPAD sensor’s reduced reliance on external light sources. Nighttime results amplify this trend, where SPAD images exhibit higher contrast, entropy, and sharpness. In brighter conditions (noon) and at longer exposures, however, the performance gap narrows, with the conventional camera matching or surpassing SPAD performance in some metrics. At these longer exposures, the conventional camera often delivers the highest metric values across the full range of illumination conditions and exposure settings. This advantage likely arises from a combination of factors. The FLIR’s higher pixel count enables it to capture finer spatial detail when sufficient photons are available. In bright, static scenes, longer exposure times allow conventional cameras (often with higher quantum efficiency and better matching sensor-optics integration) to collect significantly more photons per pixel, increasing the signal-to-noise ratio (SNR) and reducing shot noise. In contrast, the SPAD camera may reach an SNR plateau earlier due to saturation of its pixels. Note the SPAD512 model used in this evaluation did not incorporate microlenses (limiting the fill factor to $\approx 10\%$), resulting in limited photon detection efficiencies (PDEs), often below 50%. This highlights the disparities in the degree of maturity of both technologies, with conventional camera sensors benefiting from a level of technology development that single-photon sensors have yet to achieve. SPAD technology, by design, is optimized for low-light, high-speed acquisition rather than maximizing image quality in static, high photon flux conditions. Overall, these results affirm SPAD’s efficacy for low-latency, high-fidelity capture in challenging illuminations, while also highlighting context-dependent scenarios in which conventional imaging technology yet retains an advantage.

4.2. Scene segmentation

Results from running the SAM ViT-H model over a single frame captured at different illumination conditions are shown in Figure 4, with associated metrics reported in Table 3. While conclusions from a single image are necessarily limited, this frame was selected for its challenging backlit conditions and the presence of both natural (foreground rock) and artificial (lander) objects, including what can be considered an artificial horizon line in the background.

Under well-lit reference and noon scenarios, all camera configurations yield clear segmentations, with distinct object boundaries and color-coded masks. As illumination decreases, however, differences emerge. At 1 ms exposure in low light (dawn/dusk and night), the conventional camera produces underexposed images with substantial detail loss, enabling reasonable foreground segmentation but failing in the background without longer exposures or post-processing. The SPAD camera maintains robust segmentation at 8-bit during dawn/dusk, but at night, it also struggles to capture most background elements. At 4-bit depth, performance becomes less consistent, sometimes preserving object shapes, but at other times fragmenting boundaries or missing objects entirely, even when they are visually discernible. This suggests that segmentation models like SAM could benefit from training on lower bit-depth data to fully exploit the resource-efficient capabilities of SPAD cameras in photon-scarce environments.

Most conditions beyond the reference scenario result in a significant decrease in mIoU for all configurations, with the conventional camera generally achieving slightly

Table 2: Comparison between a conventional camera (FLIR Blackfly-S) and a single-photon camera (SPAD512) across multiple metrics. Values are reported in the order FLIR | SPAD and represent the average across all frames in each sequence. SPAD exposure times are given as 8-bit equivalents. Contrast and sharpness values are normalized. Highlighted entries indicate the best value for each combination of illumination condition and exposure time. The overall best values across the exposure range are highlighted in green for SPAD and in pink for FLIR. An \times denotes failed image acquisitions due to either insufficient light or excessive saturation.

Illumination	HL	Exposure (ms)	Contrast	Entropy	Sharpness	BRISQUE	MS-SSIM
Dawn/Dusk	Off	\times 0.128	\times 0.0833	\times 0.417	\times 0.006	\times 41.06	\times 0.047
		\times 0.256	\times 0.101	\times 0.344	\times 0.009	\times 40.62	\times 0.059
		1.00 1.28	0.085 0.153	0.044 0.055	0.014 0.018	51.11 39.70	0.026 0.083
		2.50 2.56	0.157 0.181	0.017 0.016	0.030 0.025	46.46 38.74	0.052 0.094
		10.00 12.8	0.215 0.229	0.00 0.063	0.029 0.073	38.08 42.59	0.153 0.033
	On	50.00 \times	0.182 \times	0.00 \times	0.0113 \times	\times \times	0.3018 \times
		\times 0.128	\times 0.177	\times 0.090	\times 0.017	\times 39.23	\times 0.029
		\times 0.256	\times 0.218	\times 0.050	\times 0.020	\times 38.02	\times 0.033
		1.00 1.28	0.137 0.289	0.009 0.00	0.016 0.02	40.03 41.23	0.029 0.058
		2.50 2.56	0.244 0.273	0.002 0.00	0.031 0.016	35.08 37.56	0.054 0.056
Noon	Off	10.00 12.80	0.296 0.146	0.00 0.022	0.027 0.024	33.19 48.16	0.066 0.022
		50.00 \times	0.181 \times	0.00 \times	0.007 \times	44.71 \times	0.055 \times
		\times 0.128	\times 0.109	\times 0.241	\times 0.009	\times 40.57	\times 0.02
		\times 0.256	\times 0.130	\times 0.163	\times 0.012	\times 39.64	\times 0.037
		1.00 1.28	0.122 0.182	0.036 0.0154	0.021 0.017	45.26 38.50	0.093 0.148
	On	2.50 2.56	0.220 0.206	0.011 0.0017	0.047 0.023	41.14 37.91	0.117 0.148
		10.00 12.80	0.279 0.209	0.00 0.059	0.043 0.041	33.63 43.37	0.146 0.010
		50.00 \times	0.188 \times	0.00 \times	0.010 \times	39.91 \times	0.192 \times
		\times 0.128	\times 0.173	\times 0.042	\times 0.018	\times 40.61	\times 0.015
		\times 0.256	\times 0.207	\times 0.014	\times 0.020	\times 38.86	\times 0.100
Night	Off	1.00 1.28	0.1493 0.249	0.005 0.00	0.022 0.017	38.82 41.08	0.087 0.096
		2.50 2.56	0.255 0.226	0.0009 0.00	0.041 0.013	35.85 38.51	0.114 0.084
		10.00 12.80	0.268 0.093	0.00 0.026	0.0271 0.032	34.92 47.71	0.087 0.009
		50.00 \times	0.144 \times	0.00 \times	0.005 \times	46.67 \times	0.038 \times
		1.28 1.28	0.003 0.036	0.00 0.171	0.00 0.019	\times 45.72	0.020 0.049
	On	2.56 2.56	0.0095 0.044	0.0006 0.261	0.0001 0.029	\times 44.67	0.070 0.040
		12.80 12.80	0.0361 0.074	0.00 0.010	0.00 0.068	46.12 \times	0.269 0.033
		128.00 128.00	0.114 0.106	0.00 0.020	0.007 0.142	36.23 40.82	0.404 0.025
		307.20 307.20	0.036 0.104	0.00 0.0200	0.00 0.135	\times 42.97	0.044 0.113
		\times 0.128	\times 0.170	\times 0.387	5 \times 0.014	\times 43.51	\times 0.005
On	\times 0.256	\times 0.220	\times 0.247	\times 0.018	\times 43.09	\times 0.007	
	1.00 1.28	0.098 0.339	0.016 0.030	0.002 0.023	43.99 42.65	0.021 0.030	
	2.50 2.56	0.207 0.353	0.006 0.0017	0.010 0.021	40.67 39.66	0.058 0.037	
	10.00 12.80	0.319 0.291	0.00 0.058	0.021 0.070	36.35 47.06	0.079 0.045	
	50.00 \times	0.231 \times	0.00 \times	0.007 \times	44.67 \times	0.112 \times	

Table 3: Segmentation metrics obtained from the results presented in Figure 4.

Illumination	Camera config.	Inference (s)	mIoU	#Masks
Reference	Conventional	95.80	1.00	35
	SPAD 8bit	99.60	1.00	23
	SPAD 4bit	79.39	0.705	11
Noon	Conventional	89.71	0.271	38
	SPAD 8bit	88.30	0.265	21
	SPAD 4bit	119.12	0.191	15
Dawn/Dusk	Conventional	105.23	0.231	29
	SPAD 8bit	97.24	0.226	27
	SPAD 4bit	100.22	0.131	21
Night	Conventional	91.94	0.122	21
	SPAD 8bit	98.27	0.265	19
	SPAD 4bit	92.48	0.3383	9

higher mIoU and detecting more masks at comparable inference times. Notably, at night, bit depth does not strictly correlate with mIoU: the 4-bit SPAD achieves a higher mIoU (0.338) than the conventional camera (0.122), indicating that performance rankings can invert in extremely low light. Inference time varies modestly across cameras and illumination conditions, with no clear link to segmentation accuracy.

While mIoU offers a convenient single-value measure, it has limitations under poor visual conditions. Low mIoU may stem from missed detections rather than poor pixel alignment, and differences in exposure or sensor characteristics can exaggerate performance gaps. For instance, at night, the conventional camera detects 21 masks with a very low mIoU, while the 4-bit SPAD detects only 9 masks but achieves a higher mIoU. This highlights the need to interpret mIoU alongside mask counts and quali-

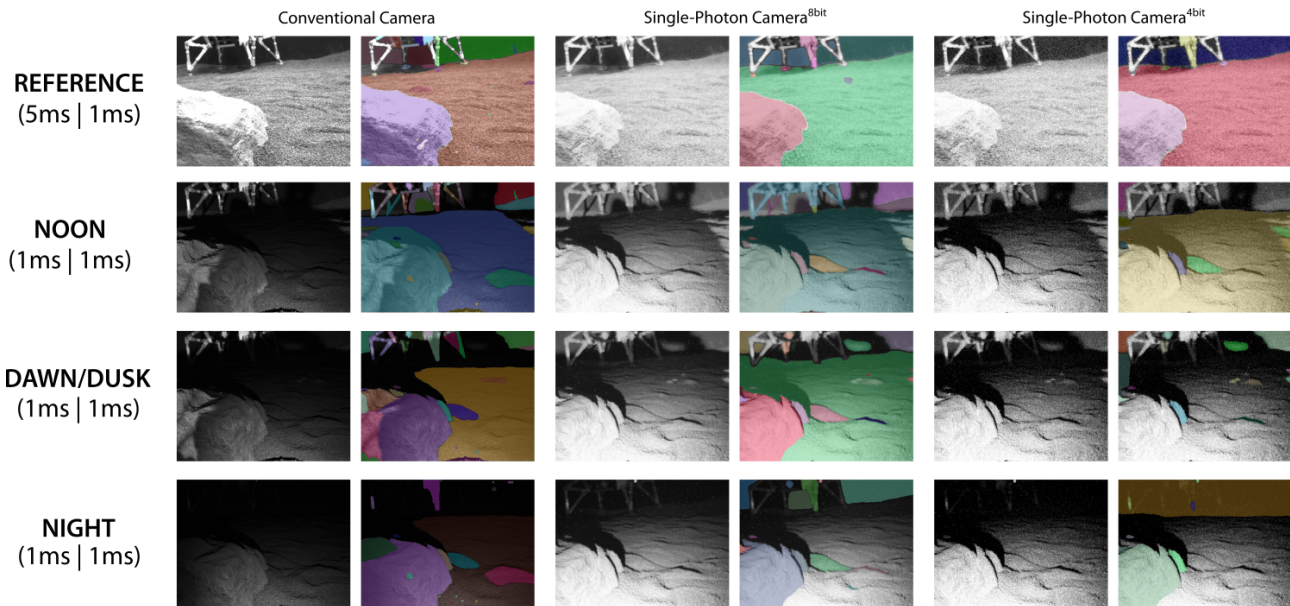


Figure 4: Segmented images resulted from running SAM [31] over conventional and single-photon images captured under different illumination scenarios. All masks are semantic-free and randomly colorized. Segmentations from 4-bit single-photon frames are also illustrated, to demonstrate the SPAD’s potential ability to reproduce comparable results even at reduced bit depths, despite models like SAM not being originally designed for low-dimensional data. Note that the equivalent exposure time at 4 bit-per-pixel would be $80\mu\text{s}$.

tative inspection for a more comprehensive assessment of segmentation performance.

4.3. Lander detection

Under well-lit conditions, both camera types have proven to capture the scene with high fidelity. To further test the extreme low-light performance of the single-photon camera, however, we wanted to evaluate segmentation results using the same image from the previous experiment, but acquired under noon conditions with the rover headlights switched off. Figure 5 shows the results of this evaluation. Unlike the conventional camera, the SPAD camera successfully resolved most of the visible lander’s structure under these conditions.

4.4. Visual Odometry

Lastly, we evaluated the performance of both camera types in a localization task. Figure 6 compares the trajectories estimated by ORB-SLAM3 using images captured with both cameras under trajectories B (parallel to sunlight) and C (toward sunlight) of the SPICE-HL3 dataset. These images were captured during dawn/dusk conditions, with the rover’s headlights on and moving at a speed of 0.5 m/s. For trajectory D (away from sunlight), neither camera recorded sufficient image features for a successful localization, resulting in the most challenging scenario. Despite the extreme visual conditions—including complex lighting and the limited field of view of both cameras (conventional camera im-

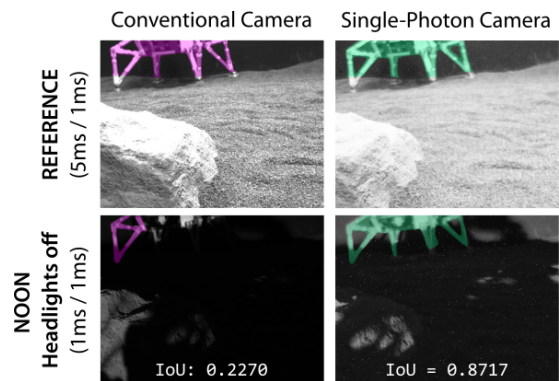


Figure 5: Results from extracting a SAM-based mask of the background lander structure on both conventional and single-photon camera images.

ages were cropped to match the aspect ratio and field of view of the SPAD)—the SPAD camera consistently produced trajectories that more closely followed the ground truth. Note that when facing direct sunlight, conventional camera images were heavily affected by lens flares and oversaturation, which prevented enough features from being identified and tracked, making localization impossible. In contrast, the SPAD camera proved more reliable under strong illumination and shadow, demonstrating robust performance in HDR lighting conditions. These results underscore the advantage of single-photon imaging for visual odometry in environments with perceptually degraded conditions. As with the SAM model used in the scene segmentation evaluation, it is important to emphasize that pipelines such as ORB-SLAM3 are ill-suited for processing low-dimensional data, which is made avail-

able directly from the SPAD raw binary stream, making it impossible to fully assess the capabilities of the SPAD in its entirety.

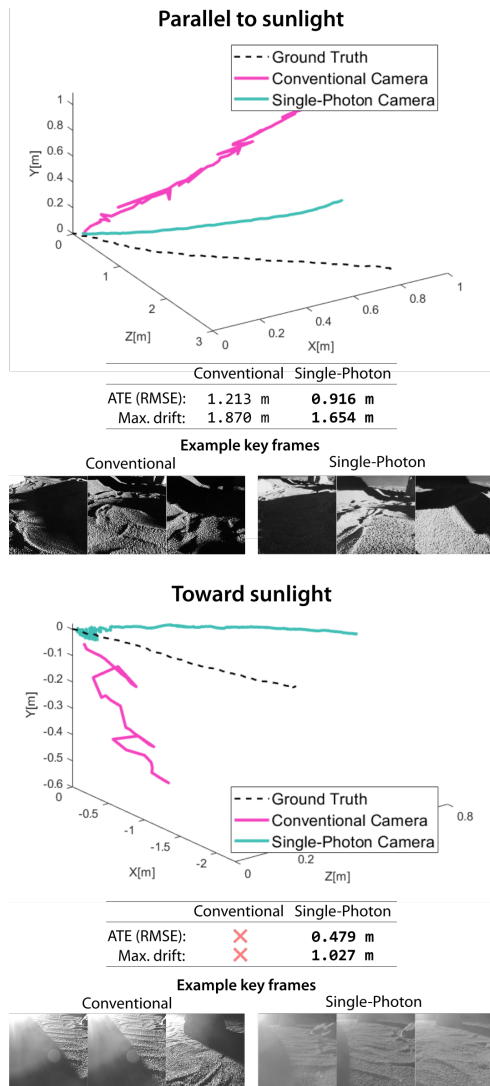


Figure 6: Results from running ORBSLAM3 over two types of trajectories: parallel to sunlight (top) and toward sunlight (bottom). Metrics are computed based on a first-pose alignment of both estimated and ground truth trajectories.

5. CONCLUSIONS

Our work confirmed that SPAD cameras demonstrate clear advantages over conventional sensors in perceptually degraded environments, particularly those affected by extremely low illuminance or demanding sensitivity to an HDR scene. Under suboptimal illumination, conventional sensors require either longer exposures or rapid exposure adaptations, both of which can exacerbate noise in images captured from fast-moving platforms, such as those moving exceedingly fast, like spacecraft in orbit. This reduced SNR limits the effectiveness of downstream computational imaging methods. However, to match the performance and efficiency of established technologies

under high photon flux conditions, further advancements are required. Recent developments, such as the integration of microlenses to improve pixel fill factor and photon detection efficiency, the addition of color filters, and the increasing sensor resolution, are promising steps toward addressing these limitations.

This evaluation is inherently constrained by the algorithms used, many of which assume 8-bit intensity data as input. To fully exploit the potential of SPADs in space and robotic applications, new algorithms must be designed to operate directly on the lower-dimensional high-throughput data SPADs produce, particularly their raw binary photon detection streams. Specialized preprocessing pipelines will also be essential for handling this data format efficiently. However, working directly with raw binary streams introduces bandwidth challenges; for many tasks, transmitting all recorded binary frames may not be neither necessary nor optimal. For example, excessively high frame rates can degrade performance in monocular visual odometry, where a minimum amount of inter-frame motion is required for a stable initialization. Adaptive frame-rate control could mitigate this issue while conserving data bandwidth.

Adaptability is, indeed, one of the two main advantages of single-photon cameras, i.e., the ability to deliver tailored, size-efficient bit-depth images at any frame rate required by the task at hand. The other is their exceptional light sensitivity and speed. An underexplored feature of SPADs for imaging under these types of illumination conditions is their potential for ultra-fast HDR imaging, achieved by combining binary streams captured at different exposure settings in near real time.

Our results show that single-photon technology can deliver robust imaging in conditions where conventional sensors often fail, highlighting SPADs' potential for autonomous robots operating in some of the most visually challenging planetary environments.

ACKNOWLEDGMENTS

This work was conducted under the project INSIGHT (PID2024-160373OB-C21) funded by the Spanish Ministry of Science, Innovation, and Universities, the Spanish State Research Agency, and the European Regional Development Fund (ERDF, UE). We extend our appreciation to Pi Imaging Technology for lending us the SPAD camera that made this study possible.

REFERENCES

1. V. Verma, M. W. Maimone, D. M. Gaines, R. Francis, T. A. Estlin, S. R. Kuhn, G. R. Rabideau, S. A. Chien, M. M. McHenry, E. J. Graser, A. L. Rankin, and E. R. Thiel, "Autonomous robotics is driving Perseverance rover's progress on Mars," *Science Robotics*, vol. 8, no. 80, p. eadi3099, 2023.
2. D. S. Bayard, D. T. Conway, R. Brockers, J. H. Delaune, L. H. Matthies, H. F. Grip, G. B. Merewether, T. L. Brown, and A. M. San Martin, "Vision-based navigation for the NASA mars helicopter," in *AIAA Scitech 2019 Forum*, p. 1411, 2019.

3. D. Rodríguez-Martínez, M. Van Winnendael, and K. Yoshida, "High-speed mobility on planetary surfaces: A technical review," *Journal of Field Robotics*, vol. 36, no. 8, pp. 1436–1455, 2019.
4. J. Trimble, "VIPER Lunar Rover Agile Mission Systems," in *17th International Conference on Space Operations (SpaceOps)*, 2023.
5. R. Soussan, J. McCaffery, S. McMichael, and M. Deans, "LuVo: Lunar Visual Odometry using Homography-based Image Feature Matching," in *IEEE International Conference on Robotics and Automation (ICRA)*, (Atlanta, USA), May 2025.
6. A. Vidal, H. Rebecq, T. Horstschaefer, and D. Scaramuzza, "Ultimate SLAM? combining events, images, and IMU for robust visual SLAM in HDR and high-speed scenarios," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 994–1001, 2018.
7. F. Mahlknecht, D. Gehrig, J. Nash, F. M. Rockenbauer, B. Morrell, J. Delaune, and D. Scaramuzza, "Exploring event camera-based odometry for planetary robots," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8651–8658, 2022.
8. V. Polizzi, R. Hewitt, J. Hidalgo-Carrió, J. Delaune, and D. Scaramuzza, "Data-efficient collaborative decentralized thermal-inertial odometry," *IEEE Robotics and Automation Letters*, pp. 1–8, 2022.
9. R. Castilla-Arquillo, A. Mandow, C. J. Pérez-del Pulgar, C. Álvarez-Llamas, J. M. Vadillo, and J. Laserna, "Thermal imagery for rover soil assessment using a multipurpose environmental chamber under simulated Mars conditions," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–12, 2023.
10. J. A. Christian and S. Cryan, "A survey of LIDAR technology and its use in spacecraft relative navigation," in *AIAA Guidance, Navigation, and Control (GNC) Conference*, p. 4641, 2013.
11. D. E. Smith, M. T. Zuber, H. V. Frey, J. B. Garvin, J. W. Head, D. O. Muhleman, G. H. Pettengill, R. J. Phillips, S. C. Solomon, H. J. Zwally, *et al.*, "Mars Orbiter Laser Altimeter: Experiment summary after the first year of global mapping of Mars," *Journal of Geophysical Research: Planets*, vol. 106, no. E10, pp. 23689–23722, 2001.
12. H. Hussmann, J. Oberst, A. Stark, and G. Steinbrügge, "The Bepi-Colombo Laser Altimeter (BELA): an instrument for geodetic investigations of Mercury," in *Planetary remote sensing and mapping*, pp. 49–61, CRC Press, 2018.
13. J. M. Leonard, M. C. Moreau, P. G. Antreasian, K. M. Getzandanner, E. Church, C. Miller, M. G. Daly, O. S. Barnouin, and D. S. Lauretta, "Cross-calibration of GNC and OLA LIDAR systems onboard OSIRIS-REx," in *Proceedings of the 44th Annual American Astronautical Society Guidance, Navigation, and Control Conference*, pp. 1585–1607, 2022.
14. M. Pyrak and J. Anderson, "Performance of Northrop Grumman's mission extension vehicle (MEV) RPO imagers at GEO," in *Autonomous systems: Sensors, processing and security for ground, air, sea and space vehicles and infrastructure 2022*, vol. 12115, pp. 64–82, 2022.
15. I. M. Antolovic, S. Burri, R. A. Hoebe, Y. Maruyama, C. Bruschi, and E. Charbon, "Photon-counting arrays for time-resolved imaging," *Sensors*, vol. 16, p. 1005, Jun 2016.
16. F. Gramuglia, E. Ripiccini, C. A. Fenoglio, M.-L. Wu, L. Paolozzi, C. Bruschi, and E. Charbon, "Sub-10 ps minimum ionizing particle detection with geiger-mode APDs," *Frontiers in Physics*, vol. 10, 2022.
17. F. Gramuglia, P. Keshavarzian, E. Kizilkan, C. Bruschi, S. S. Tan, M. Tng, E. Quek, M.-J. Lee, and E. Charbon, "Engineering breakdown probability profile for PDP and DCR optimization in a SPAD fabricated in a standard 55 nm BCD process," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 2: Optical Detectors, pp. 1–10, 2022.
18. E. R. Fossum, J. Ma, S. Masoodian, L. Anzagira, and R. Zizza, "The quanta image sensor: Every photon counts," *Sensors*, vol. 16, no. 8, p. 1260, 2016.
19. J. Ma, D. Zhang, D. Robledo, L. Anzagira, and S. Masoodian, "Ultra-high-resolution quanta image sensor with reliable photon-number-resolving and high dynamic range capabilities," *Scientific Reports*, vol. 12, no. 1, p. 13869, 2022.
20. E. Charbon, "Single-photon imaging in complementary metal oxide semiconductor processes," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 372, no. 2012, 2014.
21. K. Morimoto, A. Ardelean, M.-L. Wu, A. C. Ulku, I. M. Antolovic, C. Bruschi, and E. Charbon, "Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications," *Optica*, vol. 7, pp. 346–354, Apr 2020.
22. K. Morimoto, J. Iwata, M. Shinohara, H. Sekine, A. Abdelghafar, H. Tsuchiya, Y. Kuroda, K. Tojima, W. Endo, Y. Maehashi, *et al.*, "3.2 megapixel 3D-stacked charge focusing spad for low-light imaging and depth sensing," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 20–2, 2021.
23. V. Zickus, M.-L. Wu, K. Morimoto, V. Kapitany, A. Fatima, A. Turpin, R. Insall, J. Whitelaw, L. Machesky, C. Bruschi, F. Daniele, and E. Charbon, "Fluorescence lifetime imaging with a megapixel SPAD camera and neural network lifetime estimation," *Scientific Reports*, vol. 10, pp. 1–10, 2020.
24. S.-Y. Tsai, Y.-C. Chang, and T.-H. Sang, "SPAD LiDARs: Modeling and algorithms," in *14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, pp. 1–4, 2018.
25. Y. Liu, F. Gutierrez-Barragan, A. Ingle, M. Gupta, and A. Velten, "Single-photon camera guided extreme dynamic range imaging," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 1575–1585, 2022.
26. S. Ma, P. Mos, E. Charbon, and M. Gupta, "Burst vision using single-photon cameras," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 5375–5385, 2023.
27. L. J. Koerner, S. Gupta, A. Ingle, and M. Gupta, "Photon inhibition for energy-efficient single-photon imaging," in *European Conference on Computer Vision*, pp. 90–107, Springer, 2024.
28. M.-L. Wu, E. Ripiccini, E. Kizilkan, F. Gramuglia, P. Keshavarzian, C. A. Fenoglio, K. Morimoto, and E. Charbon, "Radiation hardness study of single-photon avalanche diode for space and high energy physics applications," *Sensors*, vol. 22, no. 8, p. 2919, 2022.
29. A. Guerrieri, D. Rodríguez-Martínez, and E. Charbon, "A dynamically reconfigurable single-board computer for high dynamic range space cameras," in *2024 IEEE 10th International Conference on Space Mission Challenges for Information Technology (SMC-IT)*, pp. 1–10, IEEE, 2024.
30. F. Yang, Y. M. Lu, L. Sbaiz, and M. Vetterli, "Bits from photons: Oversampled image acquisition using binary poisson statistics," *IEEE Transactions on image processing*, vol. 21, no. 4, pp. 1421–1436, 2011.
31. A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, *et al.*, "Segment anything," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 4015–4026, 2023.
32. C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. Montiel, and J. D. Tardós, "ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM," *IEEE transactions on robotics*, vol. 37, no. 6, pp. 1874–1890, 2021.
33. D. Rodríguez-Martínez, D. van der Meer, J. Song, A. Bera, C. Pérez-del Pulgar, and M. A. Olivares-Mendez, "SPICE-HL3: Single-photon, inertial, and stereo camera dataset for exploration of high-latitude lunar landscapes," *arXiv preprint arXiv:2506.22956*, 2025.