EventSpecPS: Photometric Stereo with Multispectral Reflectance Using an Event Camera

Jingqian Wu[1]*, Bohan Yu[2]*, Jun Hoong Chan[2],
Edmund Y. Lam[1] ✉, and Boxin Shi[2] ✉

[1]The University of Hong Kong [2]Peking University

Abstract. Surface geometry and multispectral reflectance estimation are essential tools for reliable physical scene understanding in applications such as industrial inspection, cultural heritage analysis, and remote sensing. To obtain dependable spectral and geometric information, conventional pipelines often require multiple multispectral images captured across viewpoints, or multispectral photometric stereo with repeated, sequential illuminations. These approaches are frequently bandwidth-intensive and slow to acquire due to their staged capture protocols. We propose **EventSpecPS**, an event-based multispectral photometric stereo system. It comprises a Spectrally-Angularly Multiplexed Encoding Device that encodes geometric and spectral cues into the event stream, which are then decoded by an Event-based Spectral-Angular Decoder to jointly recover surface normals and absolute multispectral reflectance. This enables single-stage, high-speed, and low-bandwidth acquisition and reconstruction, while remaining robust to cast shadows and specularities in practice. Extensive experiments on both synthetic and real data demonstrate state-of-the-art performance over existing methods. Code and data are available at `https://github.com/Wujingqian/EventSpecPS`.

Keywords: Event Camera · Photometric Stereo · Multispectral Reflectance Estimation

1 Introduction

Estimating surface normals and multispectral reflectance is a core problem in visual perception and computational imaging. Surface normals describe scene geometry [36], while multispectral reflectance characterizes wavelength-dependent material properties [4]. Recovering both is critical for applications such as industrial inspection [20], cultural heritage analysis [5], and remote sensing [2,14].

[1]Jingqian Wu and Edmund Y. Lam are with the Department of Electrical and Computer Engineering, The University of Hong Kong; [2]Bohan Yu, Jun Hoong Chan, and Boxin Shi are with State Key Laboratory of Multimedia Information Processing and National Engineering Research Center of Visual Technology, School of Computer Science, Peking University; Boxin Shi is also with PKU-AI[2] Robotics Joint Lab of Embodied AI. * Equal Contribution; ✉ Corresponding Authors.

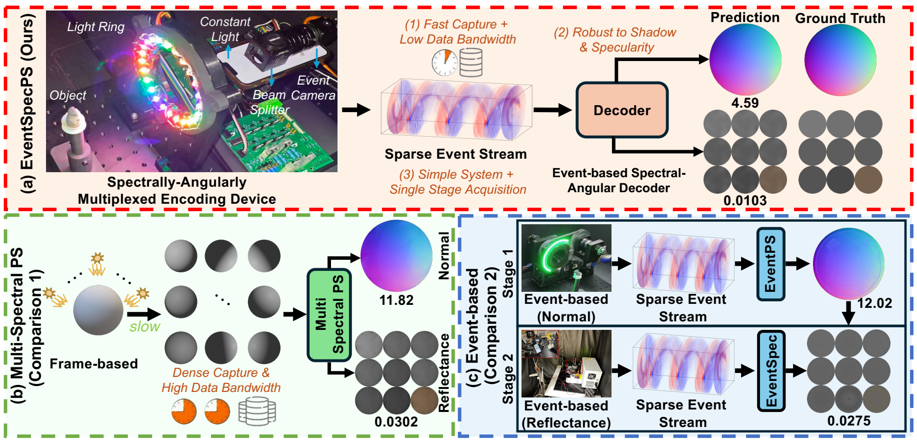

Fig. 1: The framework of EventSpecPS. (a) EventSpecPS. A single-capture event-based multispectral PS system; it has a Spectrally-Angularly Multiplexed Encoding Device to induce an event stream, which is decoded by an Event-based Spectral-Angular Decoder to jointly recover surface normals and absolute multispectral reflectance (numbers indicate errors). Compared to **(b) Frame-based multispectral PS**, EventSpecPS achieves faster, lower-bandwidth acquisition with improved robustness to specular highlights and attached shadows. It also outperforms a **(c) Two-stage Event-based alternative** that sequentially reconstructs normals and reflectance, which typically relies on heterogeneous illumination protocols and setups across stages, increases system complexity, and yields reflectance only up to an unknown scale.

Recovering geometry and multispectral reflectance has been studied through both photometric and spectral cues. Photometric Stereo (PS) infers surface normals from illumination-induced intensity variations and has been extended to both frame-based and event-based sensing [10, 27, 36, 40]. Multispectral reflectance reconstruction, on the other hand, typically relies on different sensing measurements with multiplexed illumination [4, 22, 39]. Efforts were also made to jointly formulate PS with spectral imaging, leveraging RGB-level spectral multiplexing or multiple narrow-band illuminations to couple geometry and reflectance estimation [10, 15, 22, 28].

Despite promising progress, joint recovery remains challenging. As shown in Fig. 1 **(b)**[1], frame-based joint pipelines typically rely on multispectral imaging and/or repeated captures under varying illumination, which leads to slow acquisition and high data bandwidth [10, 16, 28]. Moreover, most PS-derived formulations assume Lambertian surfaces, making reconstruction vulnerable to specular highlights and attached shadows [33], often requiring ad-hoc rejection or

[1] For reflectance visualization: the bottom-right panel visualizes the 8-channel reflectance by averaging the first 2, middle 4, and last 2-dimensional spectra and mapping it to a 3-channel RGB image, while the other patches show single-channel visualizations for channels 1–8, respectively.

additional optics for suppression [10, 16, 28]. Event-based alternatives can adopt a two-stage strategy (Fig. 1 **(c)**) to recover normals and reflectance [39, 40] with small data bandwidth and collection time; yet, such pipelines depend on different principles and illumination setups across multiple stages, resulting in increased system complexity and capturing time [27, 38, 40].

To address these challenges, we propose **EventSpecPS** (Fig. 1 **(a)**), an event-based multispectral PS system that reconstructs surface normals and *absolute* multispectral reflectance from the captured events. Specifically, we discover insights with theoretical proof showing that, under appropriate spectral-angular multiplexing conditions, there are sufficient constraints for geometry and absolute multispectral reflectance to be decoupled and determined. Guided by these insights, we co-design a *Spectrally-Angularly Multiplexed Encoding Device* that encodes geometric and spectral cues into the asynchronous event stream by rapidly switching across known spectral bands while providing diverse lighting directions. We further develop an *Event-based Spectral-Angular Decoder* to reconstruct surface normals and *absolute* multispectral reflectance from the captured events. The theoretical analysis, synthetic data, and real-world data evaluation show that EventSpecPS (i) enables fast acquisition with low data bandwidth from event streams, (ii) the reconstruction results are less affected by specular highlights and attached shadows, and (iii) supports a simple, single-stage capture system that is sufficient to solve the full task and recover *absolute* reflectance with the correct scale. Our contribution can be summarized as follows:

- We introduce **EventSpecPS**, an event-based multispectral PS system, which comprises *Spectrally-Angularly Multiplexed Encoding Device* and an *Event-based Spectral-Angular Decoder*, to jointly recover multispectral reflectance and surface normals.
- We provide theoretical proof to show our design is sufficient to decouple geometric and spectral cues, robust to specular highlights and attached shadows, and resolves the scale ambiguity.
- We conduct extensive validation using synthetic and real-world data, and compare against representative prior methods to demonstrate the feasibility and performance of EventSpecPS.

2 Related Work

2.1 Surface Normal and Multispectral Reflectance Estimation

PS estimates surface normals by observing a scene under varying illumination and has been extensively studied under both classical and learning-based settings [1, 3, 7, 9, 17, 19, 26, 36]. By modeling how surface geometry interacts with lighting, PS enables fine-grained shape recovery beyond single-image shading cues. On the other hand, Hyperspectral imaging captures wavelength-dependent radiance affected by illumination, geometry, and imaging conditions, whereas hyperspectral reflectance represents an intrinsic material property independent of lighting [8,29,32,34,41,42]. Such multispectral reflectance information is critical for applications including remote sensing, industrial inspection, and cultural heritage analysis, where accurate material interpretation is required [2,4,25].

Color or spectral-multiplexing PS methods estimate surface normals together with RGB-level reflectance, enabling dynamic capture but providing limited wavelength-resolved material information [10,28]. More recent multispectral PS studies move beyond three channels and explicitly estimate per-pixel spectral reflectance under multiple narrow-band illuminations [15,22] using hyperspectral or monochromatic cameras for joint recovery. These methods typically rely on multiple illumination conditions and frame-based hyperspectral acquisition, resulting in high data throughput and slow capture; moreover, they remain vulnerable to non-Lambertian effects such as specular highlights and attached shadows [10,16], which can degrade predictions, especially in challenging scenarios.

2.2 Event-based Photometric Analysis

Event cameras are bio-inspired sensors that asynchronously record per-pixel brightness changes with high temporal resolution, high dynamic range, and low data bandwidth consumption [11]. Benefiting from these properties, event-based vision has been actively studied across a wide range of visual tasks, including event-to-frame reconstruction [31], video interpolation [21], 3D reconstruction [37], and object detection [13]. More recently, event cameras have also attracted attention in computational photography. Event-based PS [24,27,38,40] leverages asynchronous brightness changes induced by varying illumination to recover surface normals in real time, demonstrating the feasibility of geometry estimation using event streams. Event-based hyperspectral imaging [12,38,39] explores encoding spectral information through temporally or spatially multiplexed illumination, enabling spectral reconstruction with reduced data bandwidth compared to conventional hyperspectral cameras.

Despite these promising advances, no prior event-based approach has jointly recovered surface normals and multispectral reflectance in a single system. While a two-stage combination of event-based PS [27, 40] and spectral reconstruction [38, 39] could be considered, it would require heterogeneous illumination protocols and setups across stages, leading to increased system complexity; moreover, current event-based hyperspectral imaging methods only recover reflectance up to an unknown scale.

3 Problem Formulation

3.1 Surface Normal Estimation

Assume an object is illuminated by a distant point light source whose radiance is constant over time, while the lighting direction varies and is represented by a unit vector function $\mathbf{l}(t)$. For a pixel located at image coordinate $\mathbf{x} = (x, y)$

with surface normal $\mathbf{n}_\mathbf{x}$ and diffuse albedo $a_\mathbf{x}$ (in conventional PS), the observed radiance under the Lambertian assumption is given by

$$I_\mathbf{x}^{PS}(t) = \max\left(0,\ a_\mathbf{x}\,\mathbf{n}_\mathbf{x}^\top \mathbf{l}(t)\right), \tag{1}$$

where the $\max(\cdot)$ operator accounts for attached shadows.

By observing the same surface point under multiple lighting directions, PS aims to estimate the surface normal $\mathbf{n}_\mathbf{x}$ by exploiting the variations in measured radiance induced by changes in illumination.

3.2 Multispectral Reflectance Estimation

Assume an ideal white light source with a flat spectral distribution over the visible wavelength range illuminates the scene. For a pixel at location $\mathbf{x} = (x, y)$ and wavelength λ, the multispectral image intensity I^M is modeled as

$$I_{\mathbf{x},i}^M = R_{\mathbf{x},i}\,S_i, \quad i = \{1, \ldots, N\}, \tag{2}$$

where $R_{\mathbf{x},i}$ denotes the multispectral reflectance of the surface at pixel $\mathbf{x}$, and S_i represents the spectral radiance of the illumination, which is shared across the image plane. Following common practice, we assume non-fluorescent materials and neglect inter-reflections and subsurface scattering, allowing each wavelength to be treated independently.

3.3 Event-based Normal and Multispectral Reflectance Estimation

Event cameras are bio-inspired sensors, where each pixel of an event camera independently triggers events whenever the change in log-intensity exceeds a predefined threshold [11]. Formally, let $I_\mathbf{x}(t_k)$ denote the image intensity at pixel $\mathbf{x}$ and time t_k. An event is generated at pixel $\mathbf{x}$ and timestamp t_k when the log-intensity change since the previous event exceeds a contrast threshold C. The event generation process can be expressed as an event with polarity $p_k \in \{-1, +1\}$ triggered at time t_k when

$$\log\left(I_\mathbf{x}(t_k) + \epsilon\right) - \log\left(I_\mathbf{x}(t_{k-1}) + \epsilon\right) \;\approx\; p_k\,C, \tag{3}$$

where ϵ is a small offset to avoid taking the logarithm of zero. Over a short time interval, each pixel $\mathbf{x}$ produces a sequence of events $\mathcal{E}_\mathbf{x} = \{(\mathbf{x}, p_k, t_k)\}_{k=1}^K$, where $p_k \in \{-1, +1\}$ denotes the event polarity corresponding to an increase or decrease in log-intensity.

We now unify photometric stereo, multispectral imaging, and event-based sensing into a joint formulation. For a pixel located at image coordinate $\mathbf{x} = (x, y)$, the instantaneous image intensity at time t is modeled as

$$I_\mathbf{x}(t) = \max\left(0,\ \mathbf{n}_\mathbf{x}^\top \mathbf{l}(t)\right) \sum_{i=1}^{N} D_i\,R_{\mathbf{x},i}\,S_i(t), \tag{4}$$

where $\mathbf{n}_\mathbf{x} \in \mathbb{S}^2 \subset \mathbb{R}^3$ is the unit surface normal at pixel $\mathbf{x}$, and $\mathbb{S}^2 = \{\mathbf{v} \in \mathbb{R}^3 \mid \|\mathbf{v}\|_2 = 1\}$ denotes the 2-sphere. $R_{\mathbf{x},i} \geq 0$ denotes the multispectral reflectance of the surface at pixel $\mathbf{x}$ evaluated at the i-th spectral channel λ_i, and $D_i \geq 0$ is the camera spectral response at λ_i. $\mathbf{l}(t) \in \mathbb{S}^2$ denotes the unit lighting direction at time t, determined by the active LED on the light ring, while $S_i(t) \geq 0$ represents the spectral radiance (intensity) of the illumination at channel λ_i and time t (*i.e.*, the temporal spectral modulation imposed by the light pattern).

Event cameras do not directly measure $I_\mathbf{x}(t)$, but instead respond to changes in logarithmic intensity. According to the event generation model in Eq. 3 and the "no events in attached shadow" property from [40], substituting Eq. (4) into Eq. (3), we obtain the event-based measurement constraint that couples surface geometry and multispectral reflectance:

$$\log\left(\frac{\mathbf{n}_\mathbf{x}^\top \mathbf{l}(t_k) \sum_{i=1}^{N} D_i\, R_{\mathbf{x},i}\, S_i(t_k)}{\mathbf{n}_\mathbf{x}^\top \mathbf{l}(t_{k-1}) \sum_{i=1}^{N} D_i\, R_{\mathbf{x},i}\, S_i(t_{k-1})}\right) = p_k\, C, \tag{5}$$

where $p_k \in \{-1, +1\}$ denotes the polarity of the k-th event and C is the contrast threshold of the event camera.

Our goal is to jointly estimate the surface normal field $\{\mathbf{n}_\mathbf{x}\}$ and the multispectral reflectance $\{R_{\mathbf{x},i}\}$ for all pixels $\mathbf{x}$ and spectral channels $i = \{1, \ldots, N\}$, given the event stream captured by an event camera under a time-varying spectral-angular illumination specified by $\{S_i(t), \mathbf{l}(t)\}$.

4 EventSpecPS

EventSpecPS is an event-based system for the joint recovery of surface normals and absolute multispectral reflectance from a single capture. It consists of a *Spectrally-Angularly Multiplexed Encoding Device* that encodes geometric and multispectral cues into the event stream, and an *Event-based Spectral-Angular Decoder* that decodes the events to recover surface normals and absolute per-band reflectance. However, there are **three key challenges** for naive event-based joint reconstruction. In this section, we first present the key design insights of EventSpecPS that address the challenges in event-based joint normal and multispectral reflectance recovery (Sec. 4.1). We then detail the implementation of the Spectrally-Angularly Multiplexed Encoding Device and the Event-based Spectral-Angular Decoder (Sec. 4.2).

4.1 Challenges and Design Insights

Three key challenges hinder naive event-based joint reconstruction. **First**, without additional constraints, surface normals and multispectral reflectance are inherently coupled in the event domain, making direct decoupling ill-posed.

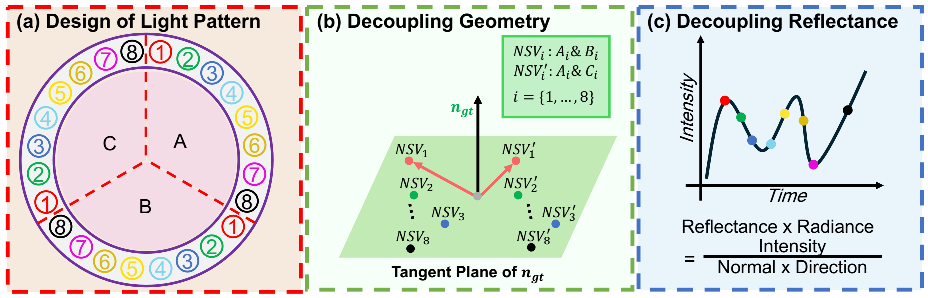

Fig. 2: EventSpecPS decouples geometry and spectral reflectance. (a) Design of Light pattern. Each spectral band is repeated three times at different angular positions (labeled A, B, and C). **(b) Decoupling geometry.** For the same spectral band, ratios between event-induced intensity changes cancel the reflectance term and yield reflectance-invariant geometric constraints; with at least two independent null space vectors (NSV), the surface normal is constrained by the intersection of the corresponding planes. **(c) Decoupling reflectance.** Once the normal is constrained, the remaining absolute-intensity information over an illumination cycle provides sufficient constraints to estimate per-band multispectral reflectance.

Second, attached shadows and specular highlights can invalidate a subset of measurements and thus reduce the effective constraints available for estimation. **Third**, since events observe only relative log-intensity changes, reflectance is, in general, identifiable only up to an unknown scale. In this subsection, we analyze these challenges and show how our design insights introduce sufficient constraints to make these problems well-conditioned.

Insight I: Decoupling Geometry and Spectral Reflectance. From Eq. (5), each event encodes a log-intensity ratio in which the surface normal $\mathbf{n_x}$ and multispectral reflectance $R_{\mathbf{x}}(\lambda)$ are multiplicatively coupled, implying that additional illumination constraints are required for decoupling. For instance, if each spectral band is activated only once (*i.e.*, without repeated observations under different lighting directions), then $\mathbf{n_x}$ and $R_{\mathbf{x}}(\lambda)$ are not identifiable: multiple normal-reflectance combinations can induce indistinguishable event responses, making direct optimization ill-conditioned.

This observation leads to a key design insight: **at least two pairs of distinct lighting directions under the same spectral channel are needed**, so that two independent intensity ratios can be formed, canceling reflectance. Each ratio yields a reflectance-invariant geometric constraint (Fig. 2 **(b)**); with two such constraints, the surface normal is determined by their intersection, up to sign. Additionally, **each spectral band must be visited at least once within an illumination cycle** (Fig. 2 **(c)**), so that the event stream provides enough intensity-change constraints beyond the reflectance-invariant ratios. Once the geometry is constrained by the ratio-based cues, the remaining intensity informa-

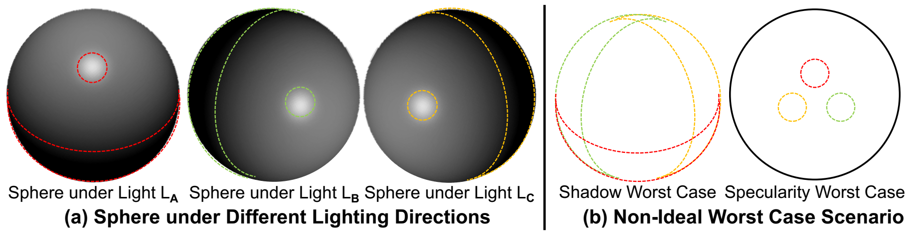

Fig. 3: Illustration of spectrum availability in a sphere. For each spectral band, **(a) Sphere under different light directions.** Our light ring provides three illuminations (L_A, L_B, L_C) at 120° azimuthal offsets. **(b) Non-ideal worst case scenario.** Attached shadows occur when $\mathbf{n}^\top \mathbf{l} \leq 0$ and manifest as a rim band that removes a contiguous set of lighting directions; even in the worst case, a substantial shadow-feasible interval remains. Specular highlights are confined to a small region near the mirror configuration, yielding only localized invalid patches. Under our setup, these two failure modes are complementary across the three repeats, so at least one observation per band remains valid for albedo estimation.

tion then becomes sufficient to estimate the multispectral reflectance. Together, these insights specify the minimal spectral-angular repetition needed to make decoupling normal and multispectral reflectance well-conditioned.

Insight II: Spectrum Availability under Shadow and Specular Effects. Attached shadows and specular highlights can invalidate individual illuminations for per-band reflectance estimation: shadows remove observations when $\mathbf{n}^\top \mathbf{l} \leq 0$, while specularities contaminate measurements when the illumination falls close to the mirror direction. In the following, we analyze these two non-idealities and detect a minimal requirement to guarantee per-band spectrum availability (at least one reliable observation for each spectral channel), even in worst-case configurations.

Attached shadow region appears as a rim band and becomes more severe near low latitude areas, as shown in Fig. 3 **(a)**; however, for any fixed light direction, an attached shadow can cover at most a hemisphere, leaving at least half of the directional space shadow-feasible. Thus, with at least three azimuthally separated repeats per spectral band, this implies that even under the worst-case shadow extent, at least one of the three observations for each band remains available (since an azimuthal offset is less than 180°), mitigating shadow-induced missing spectra. **Specular highlights** are more localized: they form compact contaminated patches rather than a global failure. We find that **the elevation angle of the lighting direction is required to be within an adequate range**, so that the specular-contaminated regions for the three repeats are separated (Fig. 3 **(b)**), and do not overlap with the shadow band, so if one observation is corrupted by a highlight, the other repeats still provide uncontaminated spectral measurements. A sufficient geometric condition and a

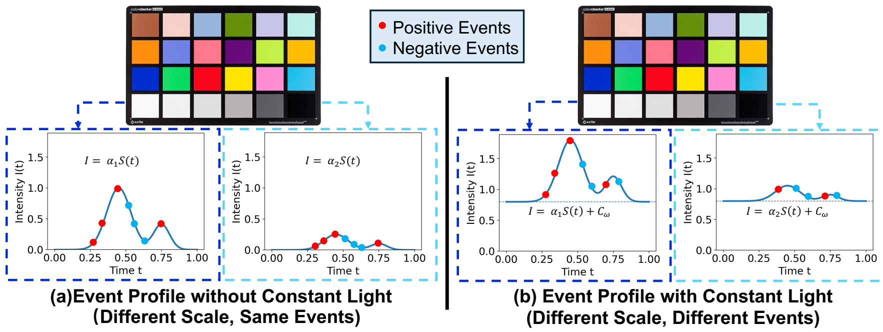

Fig. 4: Scale ambiguity. (a) Without a constant-light term, the measured intensities follow $I = \alpha S(t)$ and differ only by a multiplicative albedo scale; consequently, event observations encode identical intensity ratios for different albedos, making absolute reflectance unrecoverable. **(b)** Adding a calibrated constant-light component yields $I = \alpha S(t) + C_w$, introducing an additive offset that breaks the scale invariance; the resulting event-derived ratios become albedo-dependent, enabling closed-form recovery of absolute albedo. Colored dots indicate triggered positive and negative events.

rigorous proof of this guarantee (including the corresponding latitude range on the sphere) are provided in the supplementary material.

Insight III: Resolving Scale Ambiguity via Constant-Light Augmentation. Event cameras respond only to relative changes in log-intensity and therefore provide access to intensity ratios rather than absolute values. As a result, the reconstruction in the previous event-based method [27, 38, 40] reflects only the geometric normal, leaving an unknown scale for absolute albedo reflection. We find that **it is sufficient to recover absolute spectral reflectance by adding an illumination with a constant-light component of known radiance** C_w (similar to [39]). This additional light source can be fixed in radiance and remains active while directional illumination is switched. The theoretical detection is as follows:

With the constant-light component enabled, the observed intensity at pixel $\mathbf{x}$ under a directional illumination $\mathbf{l}$ can be written as

$$I(\mathbf{l}) = \alpha_{\mathbf{x}}\, s(\mathbf{l}) + C_w, \tag{6}$$

where $s(\mathbf{l}) = \max(0, \mathbf{n}_{\mathbf{x}}^\top \mathbf{l})$ is the geometric shading term and $\alpha_{\mathbf{x}}$ denotes the diffuse albedo. Compared to the pure Lambertian model, the presence of C_w introduces an additive offset that is independent of surface orientation.

Two known illumination directions $\mathbf{l}_1$ and $\mathbf{l}_2$, induces events whose accumulated polarity encodes the intensity ratio

$$\rho = \frac{I(\mathbf{l}_1)}{I(\mathbf{l}_2)} = \frac{\alpha_{\mathbf{x}}\, s_1 + C_w}{\alpha_{\mathbf{x}}\, s_2 + C_w}, \tag{7}$$

where $s_i = s(\mathbf{l}_i)$. Crucially, the additive constant-light term C_w breaks the scale invariance inherent to pure ratio-based observations. Solving Eq. (7) yields a closed-form expression for the absolute albedo since all other terms are known:

$$\alpha_{\mathbf{x}} = \frac{C_w(\rho - 1)}{s_1 - \rho s_2}. \tag{8}$$

A full mathematical proof can be found in the supplementary files. This analysis shows that absolute reflectance recovery does not require direct intensity measurements or additional sensors. Instead, by introducing a single calibrated constant-light component and leveraging relative intensity changes observed by the event camera, the scale ambiguity can be resolved analytically. Figure 4 illustrates this intuition: without constant-light augmentation, different albedo scales produce identical ratio responses, whereas the additive offset introduced by C_w makes absolute recovery possible.

4.2 Implementation of EventSpecPS

Spectrally-Angularly Multiplexed Encoding Device. Shown in Fig. 1(a), EventSpecPS consists of an event camera, a spectrally-angularly multiplexed light-ring illumination module, a constant-light component, and a beam splitter. The light ring contains 24 LEDs evenly distributed along the azimuth and grouped into N=8 spectral channels $\{\lambda_1, \ldots, \lambda_8\}$, where each channel appears three times at different angular positions; for channel λ_i, the corresponding lighting directions are denoted by $\mathbf{l}_i^{(1)}, \mathbf{l}_i^{(2)}$, and $\mathbf{l}_i^{(3)}$. During acquisition, exactly one LED is activated at each time instant following a predefined temporal schedule. The hardware system encodes both geometric and multispectral reflectance cues into the captured asynchronous event stream.

Event-based Spectral–Angular Decoder. Given the captured event stream and the known illumination schedule (*i.e.*, the active spectral channel, and lighting direction over time), our decoder estimates a dense surface normal map and multispectral reflectance. We parameterize the unknowns by a per-pixel normal map $\mathbf{n}$ and a rank-based multispectral reflectance model $R_{\mathbf{x},i} = \sum_{k=1}^{K} a_{\mathbf{x},k} b_{k,i}$. The reconstruction is formulated as an optimization that enforces the event-generation consistency in Eq. (5) across consecutive event pairs, together with lightweight regularization and physical constraints (non-negativity for reflectance and unit-length, front-facing normals).

5 Experiments

5.1 System and Optimization Details

Hardware details. We use a Prophesee EVK4 event sensor. To co-locate the event camera view with the constant-light path, a 50/50 beam splitter is placed in front of the sensor. To resolve the scale ambiguity inherent to event measurements, we add a constant white-light component during acquisition; in our prototype, an iPhone displaying a uniform white screen provides this constant illumination while the ring pattern is being switched. An Arduino-based controller drives the LEDs with precise timing to execute the cyclic scan pattern, including per-band repetitions and multi-round acquisition. Detailed LED specifications are provided in the supplementary material.

Optimization details. The decoder is implemented in PyTorch and optimized per scene on a single RTX 3090 GPU. We set K=8 and optimize $\mathbf{n}$, $\{a_{\mathbf{x},k}\}$, and $\{b_{k,i}\}$ using AdamW with β=$(0.5, 0.95)$, learning rates 1×10^{-2} for normals and 3×10^{-2} for reflectance parameters, for 1600 iterations per scene. We downweight specular-dominated outliers using an angular-gating weight computed from the current normal estimate and the involved light directions. We additionally use a Gaussian-blur consistency regularizer on normals and reflectance coefficients (kernel size 21, σ=5) to suppress high-frequency artifacts. We initialize a and b as positive constants with small random perturbations and initialize normals as front-facing, then enforce constraints by clamping $n_z > 0$, ℓ_2-normalizing normals, keeping reflectance non-negative, and normalizing spectral bases.

5.2 Datasets

Synthetic dataset. We build a simulation pipeline based on the renderer used in previous works [39, 40], and we extend it to support multispectral rendering (*i.e.*, outputting per-frame radiance with N spectral bands). We render objects from the Blobby dataset [18] and the Sculptures dataset [35], and we implement our light pattern in the renderer to generate paired multispectral frame sequences and the corresponding event streams. Specifically, the renderer produces multispectral frames under our light pattern, and we synthesize events from these frames using ESIM [30]. Note that frame-based baselines assume a low-rank spectral model. We therefore evaluate across synthetic datasets with increasing reflectance complexity (No. of Texture Bases = $1, 3, 8$), which cover the settings where the low-rank assumption holds and progressively departs from it. This protocol yields a fair comparison and highlights the generalization of EventSpecPS on both low- and higher-ranked multispectral reflectance.

Real-world data. To validate EventSpecPS under real sensing conditions, we capture real-world data using our hardware setup. We fabricate and capture objects with known geometry (*e.g.*, Ball, Lion, and *etc.*), for which ground-truth surface normals are available from the CAD models. To provide spectral ground truth, we additionally record patches of a ColorChecker, compute measurements of the LED spectra, and per-patch responses under white-light illumination from an optosky ATP2000H spectrometer, which enables quantitative evaluation of multispectral reflectance recovery.

5.3 Evaluation Protocol

We compare EventSpecPS against both frame-based and event-based baselines. The frame-based methods include [10, 16], while the event-based baseline is a two-stage pipeline that combines EventPS [40] with EventSpec [39]. For a fair

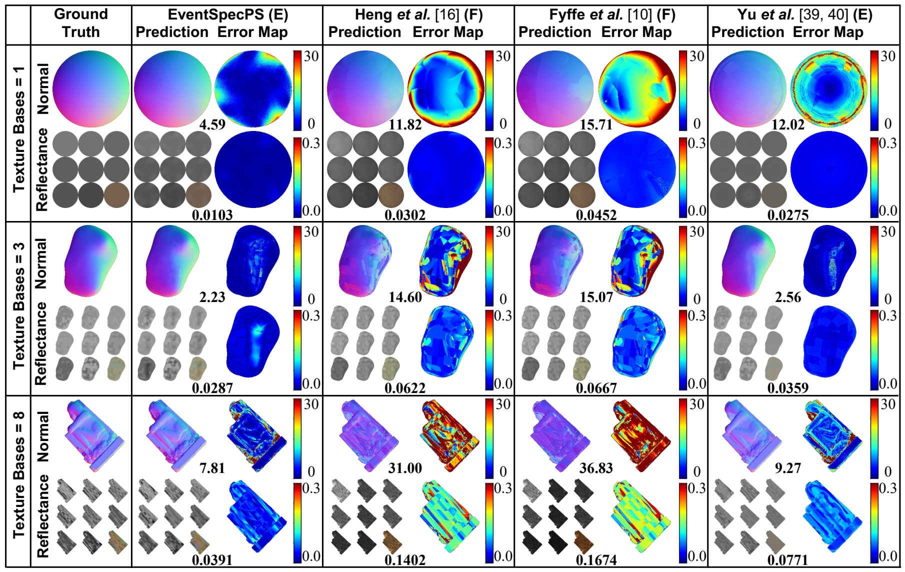

Fig. 5: Synthetic benchmark with increasing texture-bases complexity. We show predicted normals and reflectance with error maps for our method, two frame-based baselines [10, 16], and the two-stage event-based pipeline [39, 40]. Errors are in degrees (normals) and RMSE (reflectance). As complexity increases, baselines degrade with artifacts and spectral distortion, while our method remains stable.

comparison, we match the data throughput across methods as closely as possible, and we additionally report the input data volume (in MiB) used by each method. We evaluate surface normal accuracy using the mean angular error (in degrees). For multispectral reflectance, we report the mean Root Mean Square Error (RMSE) [6] and Spectral Angle Mapper (SAM) [23]. On real data, it is difficult to reproduce every baseline's hardware configuration; therefore, all quantitative comparisons against baselines are conducted on the synthetic benchmark. Real-world experiments are used to validate our system under practical non-idealities. Since it is difficult to obtain both accurate ground-truth normals and accurate ground-truth spectral reflectance for a random object with a complex texture from complex real materials, we report normal errors on fabricated objects with known geometry (Ball, Lion); for reflectance, we report RMSE on an X-Rite ColorChecker chart.

5.4 Evaluation on Synthetic Data

As shown in Fig. 5, our method consistently produces accurate normals and faithful spectral reflectance across all complexity levels. In contrast, frame-based methods exhibit increasing geometric distortion and spectral artifacts as the texture bases grow. The two-stage event-based pipeline improves upon frame-based

Table 1: Synthetic benchmark on Blobs and Sculpture datasets with different texture bases complexity. We report angular error for normal, SAM, and RMSE for reflectance. Data denotes the average input per scene. Metrics are computed on the valid-pixel mask, and reflectance is globally scaled to the ground truth for non-absolute methods.

			Normal	Reflectance		Data
Bases	Dataset	Method	MAE↓	SAM↓	RMSE↓	Avg. MiB
1	Blobs	Event-based [39, 40]	9.61	9.10	0.0302	13.31
		Frame-based (Heng *et al.*) [16]	16.01	7.58	0.0402	16.00
		Frame-based (Fyffe *et al.*) [10]	20.62	13.49	0.0591	16.00
		EventSpecPS (Ours)	**5.89**	**2.65**	**0.0164**	**12.66**
	Sculptures	Event-based [39, 40]	18.39	10.55	0.0383	7.53
		Frame-based (Heng *et al.*) [16]	21.96	15.28	0.0755	16.00
		Frame-based (Fyffe *et al.*) [10]	27.41	21.07	0.1056	16.00
		EventSpecPS (Ours)	**12.22**	**3.11**	**0.0171**	**7.26**
3	Blobs	Event-based [39, 40]	9.62	10.52	0.0633	13.31
		Frame-based (Heng *et al.*) [16]	22.65	19.71	0.0886	16.00
		Frame-based (Fyffe *et al.*) [10]	27.58	25.21	0.1058	16.00
		EventSpecPS (Ours)	**8.16**	**5.78**	**0.0394**	**12.80**
	Sculptures	Event-based [39, 40]	18.38	10.13	0.0576	7.05
		Frame-based (Heng *et al.*) [16]	28.44	26.50	0.1225	16.00
		Frame-based (Fyffe *et al.*) [10]	33.06	32.53	0.1428	16.00
		EventSpecPS (Ours)	**13.5**	**7.42**	**0.0452**	**7.16**
8	Blobs	Event-based [39, 40]	9.61	10.36	0.0663	13.31
		Frame-based (Heng *et al.*) [16]	24.07	22.05	0.1017	16.00
		Frame-based (Fyffe *et al.*) [10]	28.73	27.28	0.1182	16.00
		EventSpecPS (Ours)	**9.38**	**7.41**	**0.0471**	**13.17**
	Sculptures	Event-based [39, 40]	18.41	10.61	0.0633	6.93
		Frame-based (Heng *et al.*) [16]	29.25	27.99	0.1289	16.00
		Frame-based (Fyffe *et al.*) [10]	34.02	34.14	0.1527	16.00
		EventSpecPS (Ours)	**14.6**	**9.09**	**0.0541**	**7.00**

approaches but still suffers from accumulated errors. Not to mention, in practice, the two entirely different systems are difficult to integrate. Table 1 confirms these observations quantitatively. Our method achieves the lowest normal angular error and reflectance errors in all settings while requiring comparable or lower data throughput than event-based baselines. Notably, performance gaps of other methods widen as spectral complexity increases (Texture Bases=8), highlighting the advantage of jointly modeling geometry and reflectance.

The texture bases study is important for fairness and generalization. Frame-based reflectance solvers assume a low-rank spectral model; directly applying them to higher-dimensional spectra leads to degraded performance. By evaluating across both low- and high-bases settings, we demonstrate both a fair

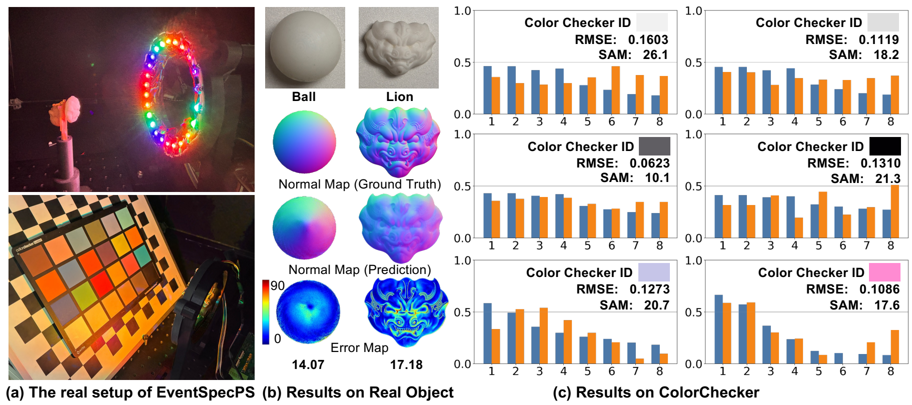

Fig. 6: Real-world evaluation of EventSpecPS. (a) The capturing process of EventSpecPS under various multispectral lights, **(b) Quantitative normal reconstruction** on fabricated objects with known CAD geometry; we show predicted normals and the corresponding angular error maps, **(c) Quantitative multispectral reflectance results on a ColorChecker:** per-patch recovered spectra across the 8 bands are compared to the reference, where the vertical axis shows the SAM and the horizontal shows the multispectral band. Blue and orange bars denote ground-truth and predictions, respectively.

comparison in favorable regimes and the robustness of our method under more challenging spectral conditions.

5.5 Evaluation on Real-world Data

Simultaneous Surface Normal and Multispectral Reflectance Estimation. We evaluate **EventSpecPS** under practical sensing non-idealities, including sensor noise, LED transient responses, and non-Lambertian effects (attached shadows and specular highlights), using data acquired with the proposed multiplexed ring pattern. As shown in Fig. 6, our real-world experiments validate that EventSpecPS produces accurate surface normals from event measurements even in the presence of attached shadows and specular highlights. For multispectral reflectance, we compare the GT spectral reflectance of patches on the X-Rite ColorChecker chart with the recovered per-band reflectance; we report RMSE and SAM over chart patches. Overall, the results demonstrate that EventSpecPS yields stable normal reconstruction and consistent per-band reflectance recovery under real sensing non-idealities.

Multispectral Imaging Capability. To further validate the multispectral imaging capability of EventSpecPS beyond the ColorChecker evaluation, we additionally capture a real colorful handwritten object on A4 paper. Fig. 7 shows the recovered reflectance across all 8 spectral channels, together with a 3-channel visualization formed by channels 0/5/3 and an iPhone reference image. While

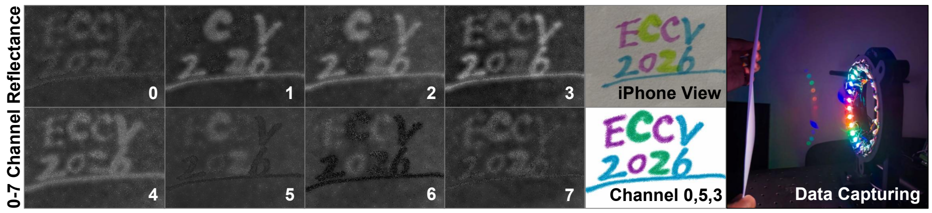

Fig. 7: We capture a colorful handwritten object and visualize the recovered reflectance across all 8 spectral channels (channels 0–7). We also show a 3-channel visualization composed from channels 0/5/3, together with an iPhone reference image. The different responses across spectral channels demonstrate that EventSpecPS recovers meaningful wavelength-dependent reflectance on a real colorful object.

this example is mainly qualitative, it provides a direct visual confirmation that different colors produce distinct responses across the recovered spectral bands. We observe that different colored strokes exhibit clearly different responses across the 8 spectral channels, indicating that the recovered reflectance preserves meaningful wavelength-dependent material/color variation. The 3-channel visualization is also consistent with the overall color layout in the iPhone image, further supporting that EventSpecPS can recover spatially coherent multispectral reflectance in real-world colorful scenes.

6 Conclusion

We presented **EventSpecPS**, an event-based multispectral photometric stereo system that recovers surface normals and *absolute* multispectral reflectance from a single capture. By combining a Spectrally–Angularly Multiplexed Encoding Device with an Event-based Spectral–Angular Decoder, EventSpecPS encodes geometric and spectral cues into the event stream and decodes them for joint reconstruction, enabling fast, low-bandwidth acquisition with improved robustness to cast shadows and specular highlights. Experiments on both synthetic and real data demonstrate state-of-the-art accuracy and consistent qualitative improvements over representative baselines.

7 Acknowledgement.

This work was supported by National Natural Science Foundation of China (Grant No. 624B2006), the Research Grants Council of Hong Kong (GRF 17201822), the Theme-based Research Scheme (T45-701/22-R), Beijing Natural Science Foundation (Grant No. L233024), Beijing Municipal Science & Technology Commission, Administrative Commission of Zhongguancun Science Park (Grant No. Z241100003524012), and Beijing High Innovation Plan (Contract No. 202604841443).

References

1. Ackermann, J., Goesele, M., et al.: A survey of photometric stereo techniques. Foundations and Trends in Computer Graphics and Vision **9**(3-4), 149–254 (2015) 3
2. Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., Sousa, J.J.: Hyperspectral imaging: A review on uav-based sensors, data processing and applications for agriculture and forestry. Remote sensing **9**(11), 1110 (2017) 1, 4
3. Bae, G., Davison, A.J.: Rethinking inductive biases for surface normal estimation. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2024) 3
4. Chang, C.I.: Hyperspectral imaging: techniques for spectral detection and classification, vol. 1. Springer Science & Business Media (2003) 1, 2, 4
5. Delaney, J.K., Zeibel, J.G., Thoury, M., Littleton, R., Palmer, M., Morales, K.M., de La Rie, E.R., Hoenigswald, A.: Visible and infrared imaging spectroscopy of picasso's harlequin musician: mapping and identification of artist materials in situ. Applied spectroscopy **64**(6), 584–594 (2010) 1
6. Dennison, P.E., Halligan, K.Q., Roberts, D.A.: A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. Remote Sensing of Environment **93**(3), 359–367 (2004) 12
7. Du, H., Yan, X., Wang, J., Xie, D., Pu, S.: Rethinking the approximation error in 3D surface fitting for point cloud normal estimation. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2023) 3
8. Enyo, Y., Nishino, K.: Diffusion reflectance map: Single-image stochastic inverse rendering of illumination and reflectance. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2024) 4
9. Fu, X., Yin, W., Hu, M., Wang, K., Ma, Y., Tan, P., Shen, S., Lin, D., Long, X.: Geowizard: Unleashing the diffusion priors for 3D geometry estimation from a single image. In: Proc. of European Conference on Computer Vision (2024) 3
10. Fyffe, G.: Single-shot photometric stereo by spectral multiplexing. In: Proc. of ACM SIGGRAPH Asia (2010) 2, 3, 4, 11, 12, 13
11. Gallego, G., Delbrück, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A.J., Conradt, J., Daniilidis, K., et al.: Event-based vision: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(1), 154–180 (2020) 4, 5
12. Geckeler, C., Neugebauer, N., Muglikar, M., Scaramuzza, D., Mintchev, S.: Event spectroscopy: Event-based multispectral and depth sensing using structured light. IEEE Robotics and Automation Letters **11**(3), 2658–2665 (2026) 4
13. Gehrig, D., Scaramuzza, D.: Low-latency automotive vision with event cameras. Nature **629**(8014), 1034–1040 (2024) 4
14. Goetz, A.F., Vane, G., Solomon, J.E., Rock, B.N.: Imaging spectrometry for earth remote sensing. Science **228**(4704), 1147–1153 (1985) 1
15. Guo, H., Okura, F., Shi, B., Funatomi, T., Mukaigawa, Y., Matsushita, Y.: Multispectral photometric stereo for spatially-varying spectral reflectances: A well posed problem? In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2021) 2, 4
16. Guo, H., Okura, F., Shi, B., Funatomi, T., Mukaigawa, Y., Matsushita, Y.: Multispectral photometric stereo for spatially-varying spectral reflectances. International Journal of Computer Vision **130**(9), 2166–2183 (2022) 2, 3, 4, 11, 12, 13
17. Ikehata, S., Asano, Y.: Physics-free spectrally multiplexed photometric stereo under unknown spectral composition. arXiv preprint arXiv:2410.20716 (2024) 3

18. Johnson, M.K., Adelson, E.H.: Shape estimation in natural illumination. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2011) 11
19. Ju, Y., Lam, K.M., Xie, W., Zhou, H., Dong, J., Shi, B.: Deep learning methods for calibrated photometric stereo and beyond. IEEE Transactions on Pattern Analysis and Machine Intelligence **46**(11), 7154–7172 (2024) 3
20. Khan, M.J., Khan, H.S., Yousaf, A., Khurshid, K., Abbas, A.: Modern trends in hyperspectral image analysis: A review. IEEE Access **6**, 14118–14129 (2018) 1
21. Kim, T., Chae, Y., Jang, H.K., Yoon, K.J.: Event-based video frame interpolation with cross-modal asymmetric bidirectional motion fields. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2023) 4
22. Kitahara, M., Okabe, T., Fuchs, C., Lensch, H.P.: Simultaneous estimation of spectral reflectance and normal from a small number of images. In: International Conference on Computer Vision Theory and Applications (2015) 2, 4
23. Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A., Barloon, P., Goetz, A.F.: The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. Remote sensing of environment **44**(2-3), 145–163 (1993) 12
24. Li, J., Zhang, Y., Han, H., Ji, X.: Active event-based stereo vision. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2025) 4
25. Li, J., Pei, Y., Zhao, S., Xiao, R., Sang, X., Zhang, C.: A review of remote sensing for environmental monitoring in china. Remote Sensing **12**(7), 1130 (2020) 4
26. Li, Q., Feng, H., Gong, X., Liu, Y.S.: Learning normals of noisy points by local gradient-aware surface filtering. In: Proc. of IEEE/CVF International Conference on Computer Vision (2025) 3
27. Liang, J., Yu, B., Yang, S., Zhuang, H., Ren, J., Duan, P., Shi, B.: EventUPS: Uncalibrated photometric stereo using an event camera. In: Proc. of IEEE/CVF International Conference on Computer Vision (2025) 2, 3, 4, 9
28. Miyazaki, D., Onishi, Y., Hiura, S.: Color photometric stereo using multi-band camera constrained by median filter and occluding boundary. Journal of Imaging **5**(7), 64 (2019) 2, 3, 4
29. Peng, S., Guo, J., Chen, X., Liu, Y., Yang, D., Bao, H., Zhou, X.: Learning diffusion priors for inverse rendering under unknown illumination. IEEE Transactions on Pattern Analysis and Machine Intelligence (2026) 4
30. Rebecq, H., Gehrig, D., Scaramuzza, D.: ESIM: an open event camera simulator. In: Conference on robot learning. PMLR (2018) 11
31. Rebecq, H., Ranftl, R., Koltun, V., Scaramuzza, D.: High speed and high dynamic range video with an event camera. IEEE Transactions on Pattern Analysis and Machine Intelligence **43**(6), 1964–1980 (2019) 4
32. Salesin, K., Seyb, D., Friday, S., Jarosz, W.: Diy hyperspectral imaging via polarization-induced spectral filters. In: Proc. of International Conference on Computational Photography (2022) 4
33. Shi, B., Wu, Z., Mo, Z., Duan, D., Yeung, S.K., Tan, P.: A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2016) 2
34. Tiwari, A., Ikehata, S., Raman, S.: MERLiN: Single-shot material estimation and relighting for photometric stereo. In: Proc. of European Conference on Computer Vision (2024) 4
35. Wiles, O., Zisserman, A.: SilNet: Single-and multi-view reconstruction by learning from silhouettes. arXiv preprint arXiv:1711.07888 (2017) 11
36. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. Optical engineering **19**(1), 139–144 (1980) 1, 2, 3

37. Wu, J., Zhu, S., Wang, C., Lam, E.Y.: Ev-GS: Event-based gaussian splatting for efficient and accurate radiance field rendering. In: IEEE International Workshop on Machine Learning for Signal Processing (2024) 4
38. Yu, B., Han, J., Shi, B., Sato, I.: EventPSR: Surface normal and reflectance estimation from photometric stereo using an event camera. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2025) 3, 4, 9
39. Yu, B., Liang, J., Wang, Z., Fan, B., Subpa-asa, A., Shi, B., Sato, I.: Active hyperspectral imaging using an event camera. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2025) 2, 3, 4, 9, 11, 12, 13
40. Yu, B., Ren, J., Han, J., Wang, F., Liang, J., Shi, B.: EventPS: Real-time photometric stereo using an event camera. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2024) 2, 3, 4, 6, 9, 11, 12, 13
41. Zeng, H., Cao, J., Zhang, K., Chen, Y., Luong, H., Philips, W.: Unmixing diffusion for self-supervised hyperspectral image denoising. In: Proc. of IEEE/CVF Computer Vision and Pattern Recognition (2024) 4
42. Zhang, J., Su, R., Fu, Q., Ren, W., Heide, F., Nie, Y.: A survey on computational spectral reconstruction methods from RGB to hyperspectral imaging. Scientific reports **12**(1), 11905 (2022) 4