北京大学计算机学院
School of Computer Science

Camera Intelligence
A Computational Photography Lab @ PKU
http://camera.pku.edu.cn

High-speed Photometric Analysis using the Neuromorphic Camera

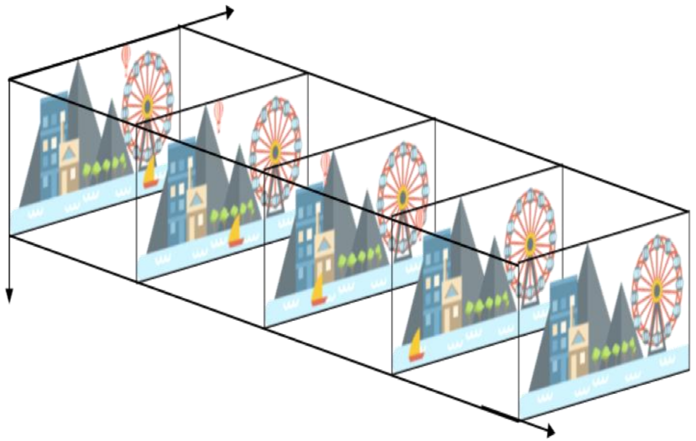
Boxin Shi
shiboxin@pku.edu.cn

2024.06.27 @ SNL2024, Tokyo, Japan

* Courtesy of APSIPA Distinguished Lecturer Program

Frame-based Cameras

The past 60 years of research have been devoted to frame-based cameras … but they are not good enough!



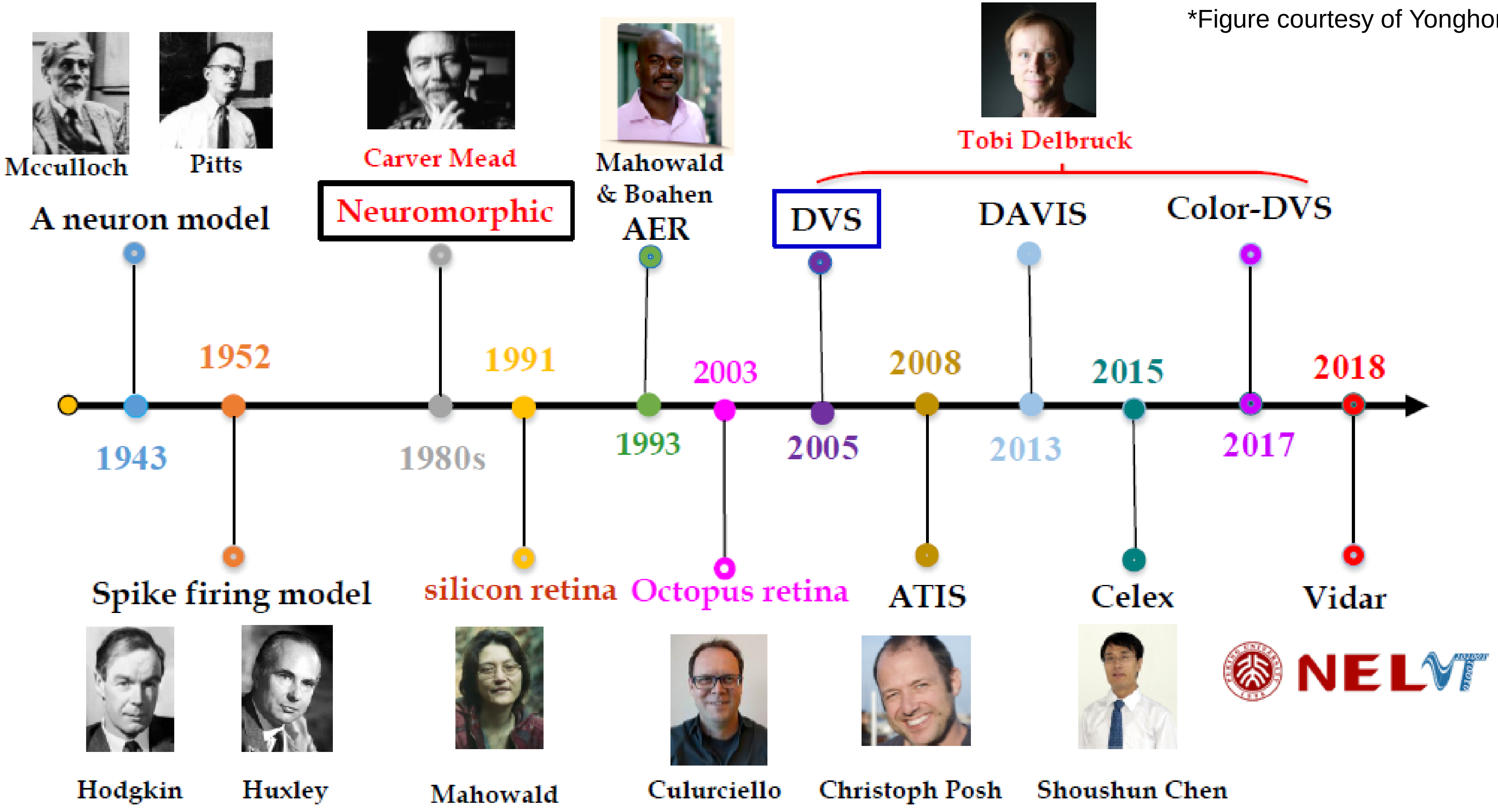
Redundant

High Latency

Low Dynamic Range

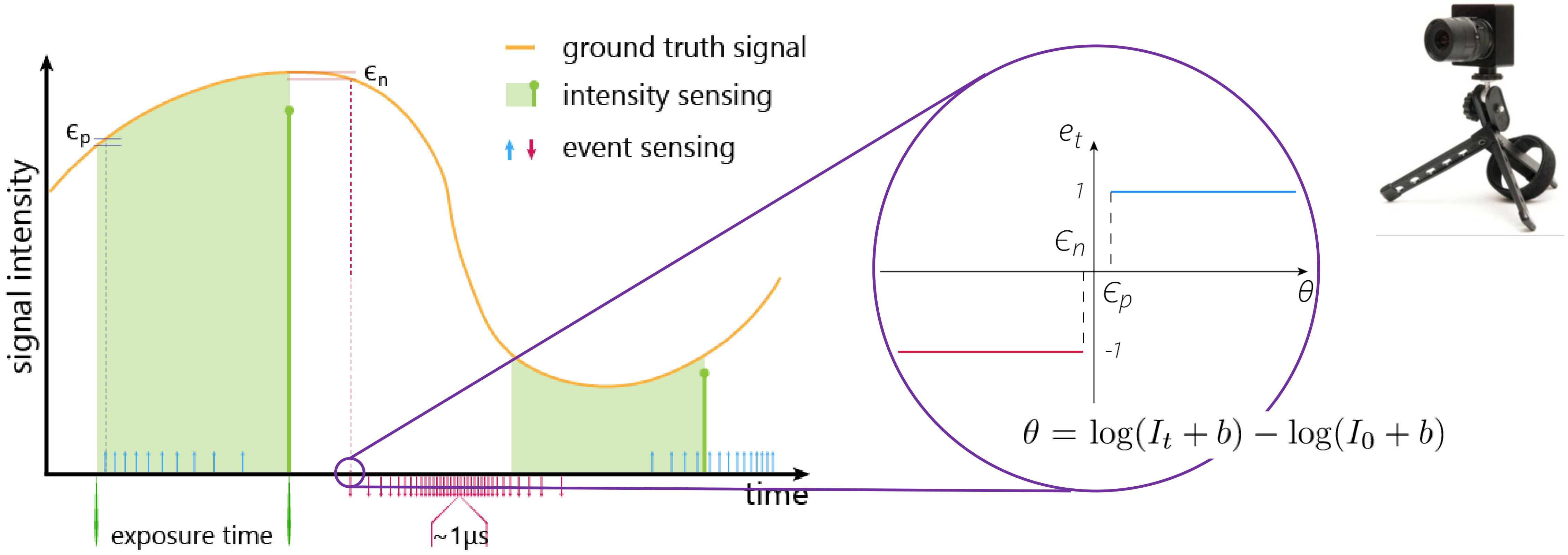
Neuromorphic Vision

*Figure courtesy of Yonghong Tian



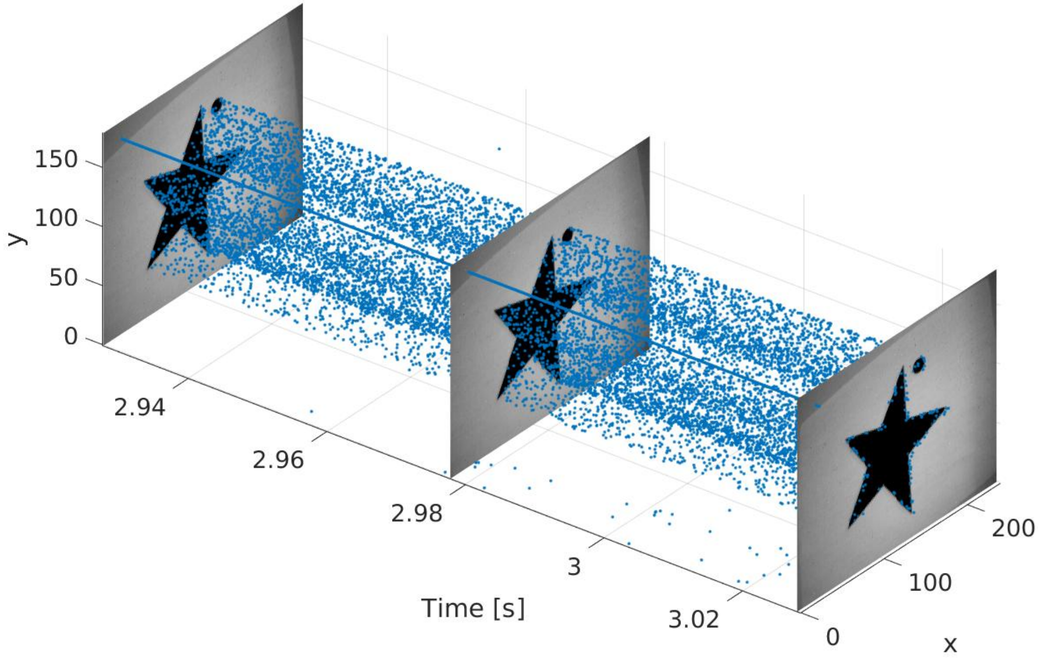
Neuromorphic Camera Example: Event Camera

- Events are generated any time a single pixel sees a change in brightness larger than the threshold



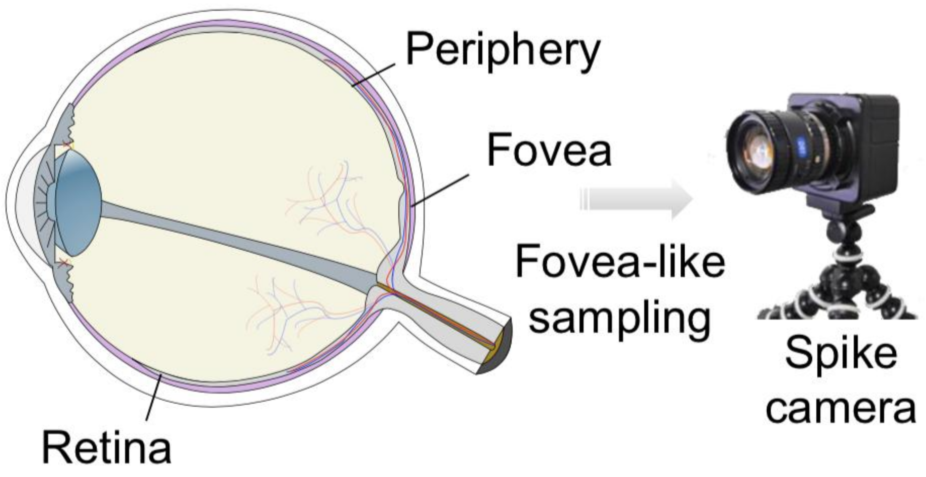
Neuromorphic Camera Example: Event Camera

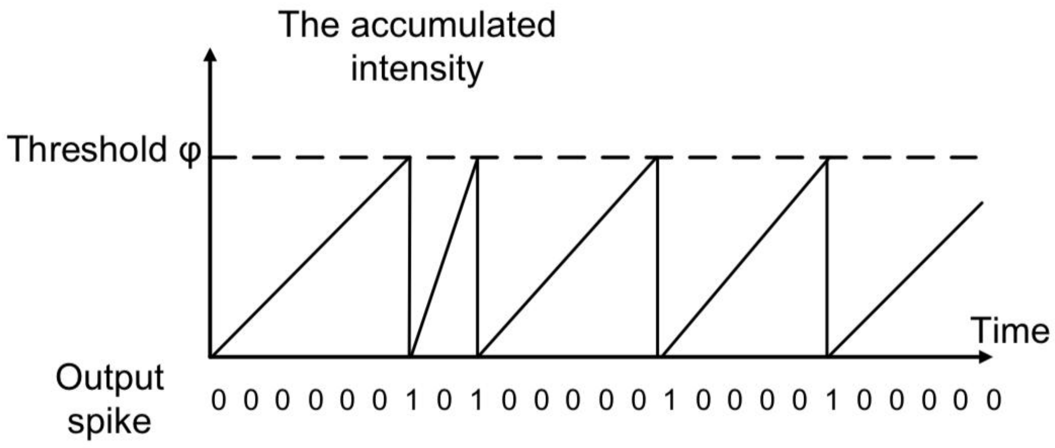
- Event camera has novel sensor that measures only motion in the scene. First commercialized in 2008 under the name of Dynamic Vision Sensor (DVS)



*Video courtesy of Elias Mueggler

Neuromorphic Camera Example: Spike Camera





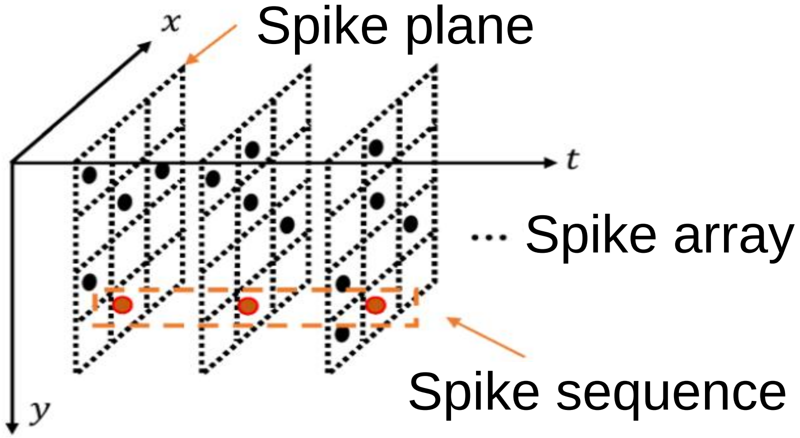
$$\int I dt \geq \phi$$

[Zhu et al., CVPR'20]

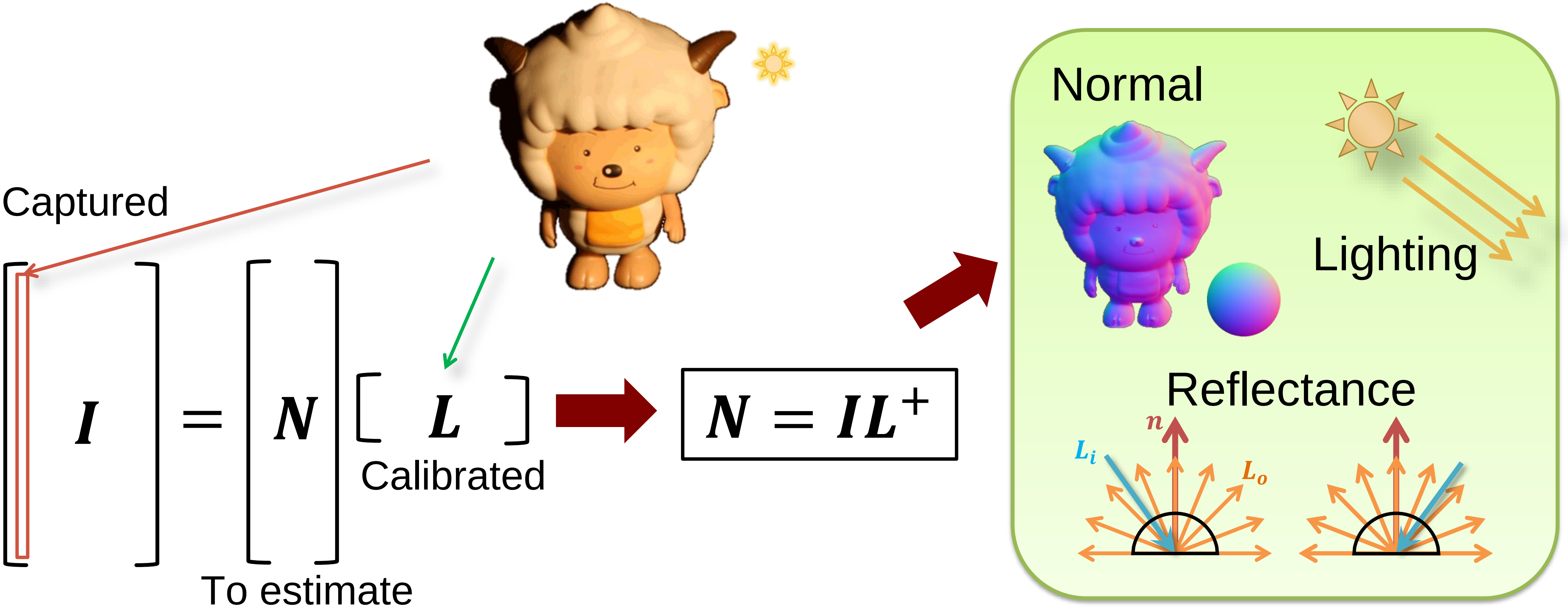
- For a pixel, the light intensity is accumulated, if the accumulated intensity reaches the dispatch threshold φ, a spike is fired and the accumulator is reset

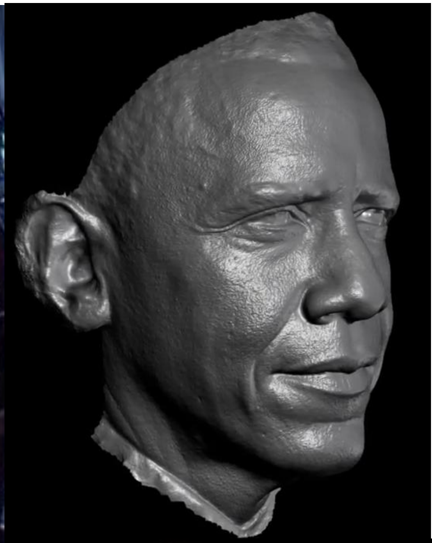
Neuromorphic Camera Example: Spike Camera

- The first spike camera is designed by **Peking university in 2018**, with spatial resolution of 400x250 and temporal resolution of 40KHz



Photometric Methods in Computer Vision

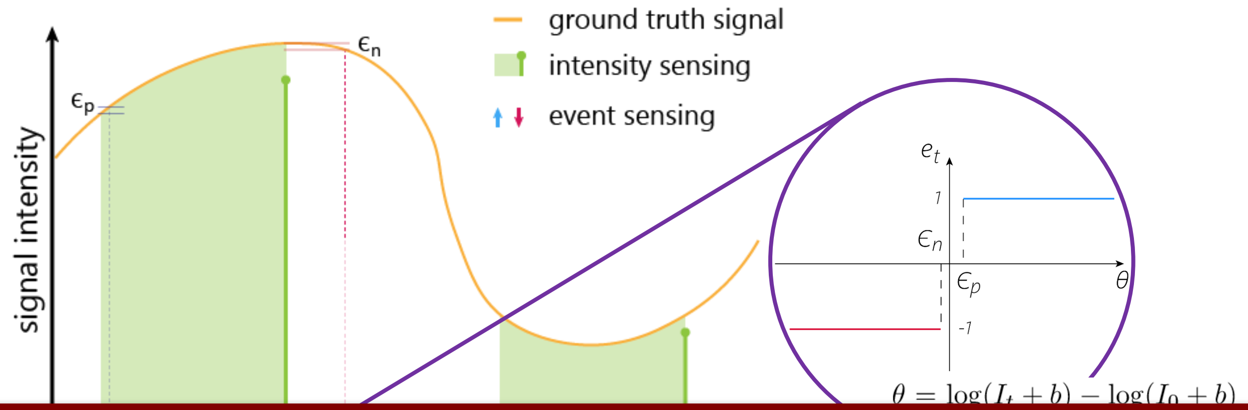




3D Scanning the President of the United States

[Debevec et al. 15]

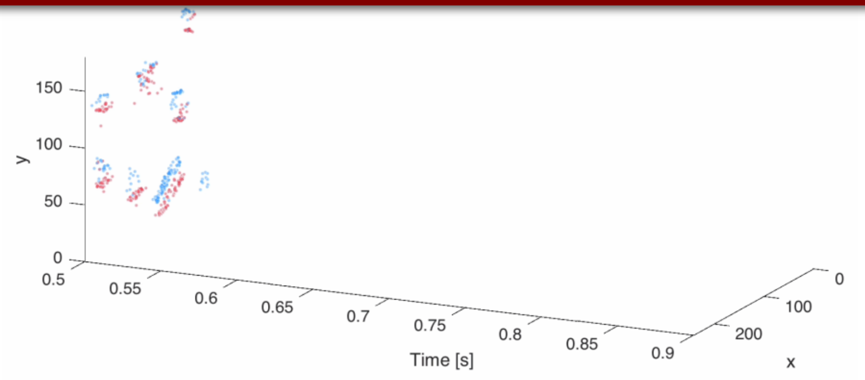
Event Meets Photometry



- Events are generated any time a single pixel sees a change in brightness larger than the threshold —

The ideal trigger model of event signal:

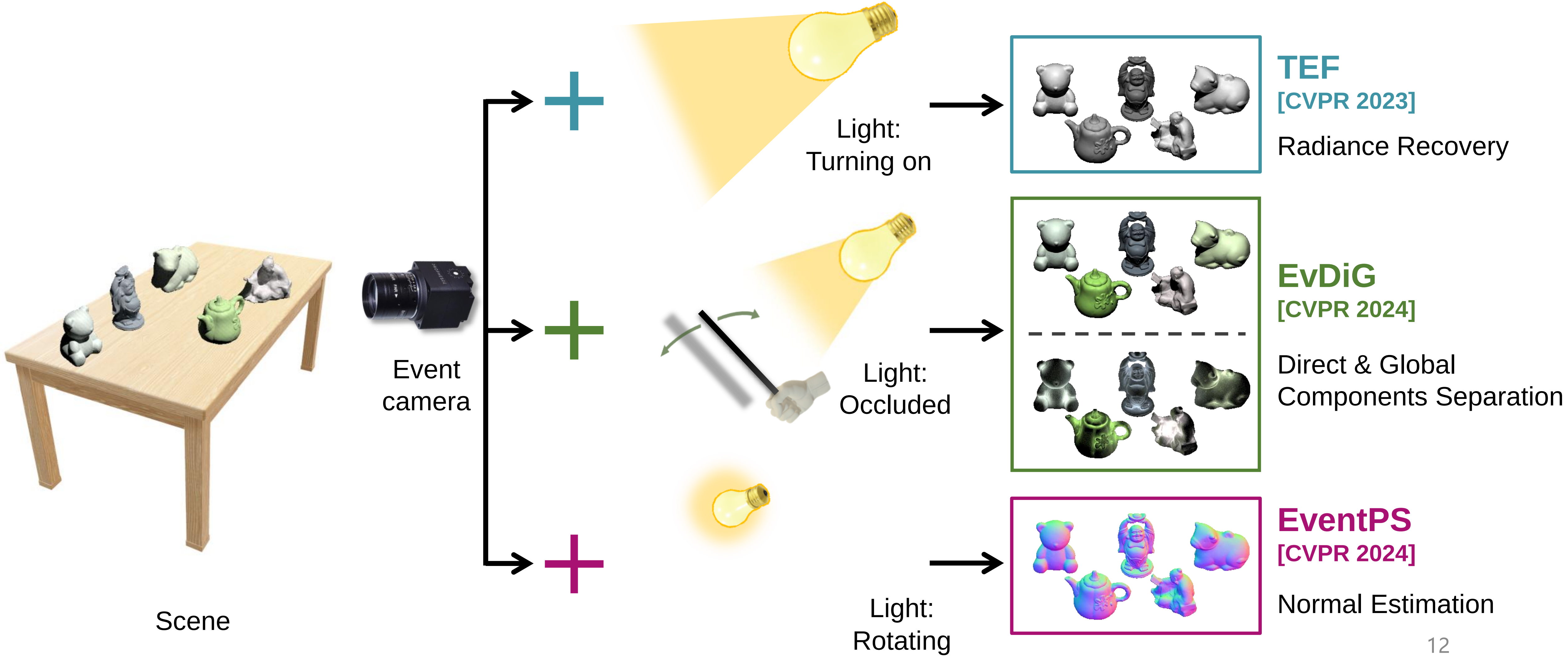
But applying such sparse/differential signals to photometric problems is not intuitive. How can an event camera benefit photometric vision?



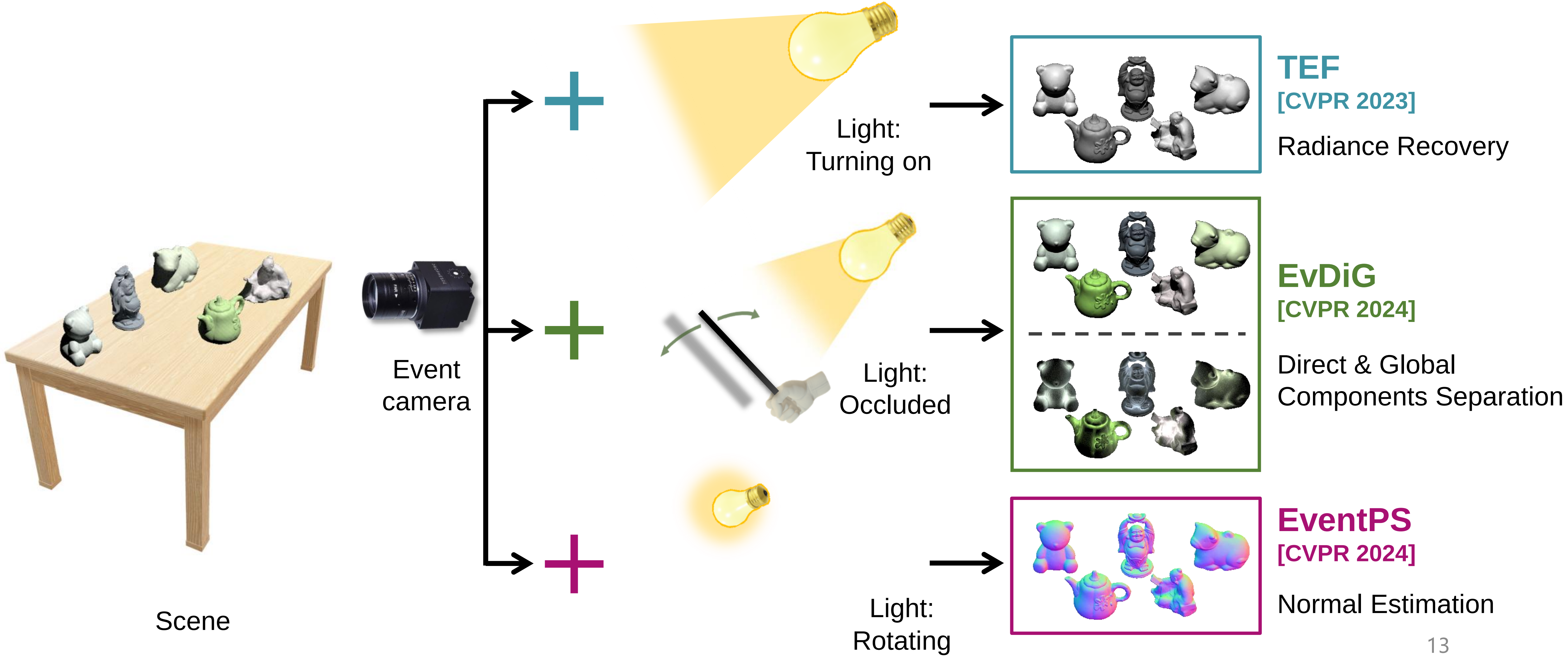
Advantages

- **High** temporal resolution (1 MHz)
- **HDR** (>120dB)
- **Low** power (10mW)

Event Photometry



Event Photometry



CVPR 2023

High-fidelity Event-Radiance Recovery via Transient Event Frequency

Jin Han[1,2] Yuta Asano[2] Boxin Shi[3,4] Yinqiang Zheng[1] Imari Sato[1,2]

[1] The University of Tokyo [2] National Institute of Informatics

[3] National Key Laboratory for Multimedia Information Processing, School of Computer Science, Peking University

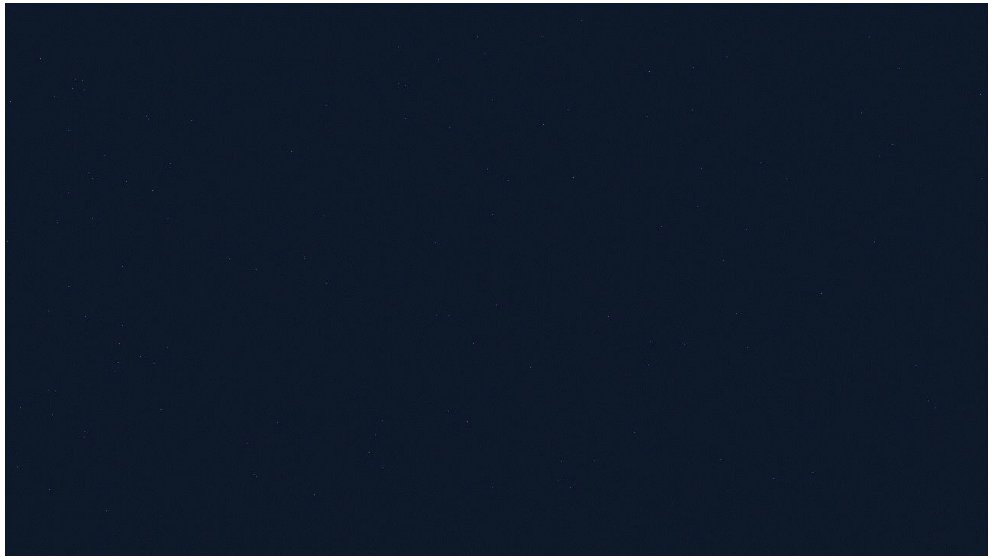
[4] National Engineering Research Center of Visual Technology, School of Computer Science, Peking University

Frame-based Camera Observation

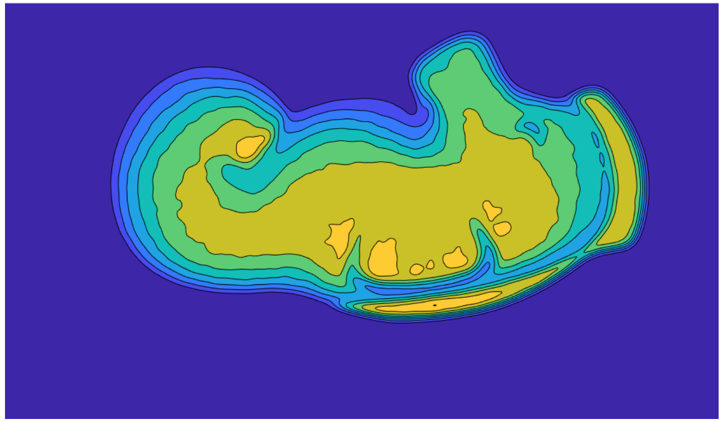
When turning on the light…


What can we get from such a short period?

Event Camera Observation

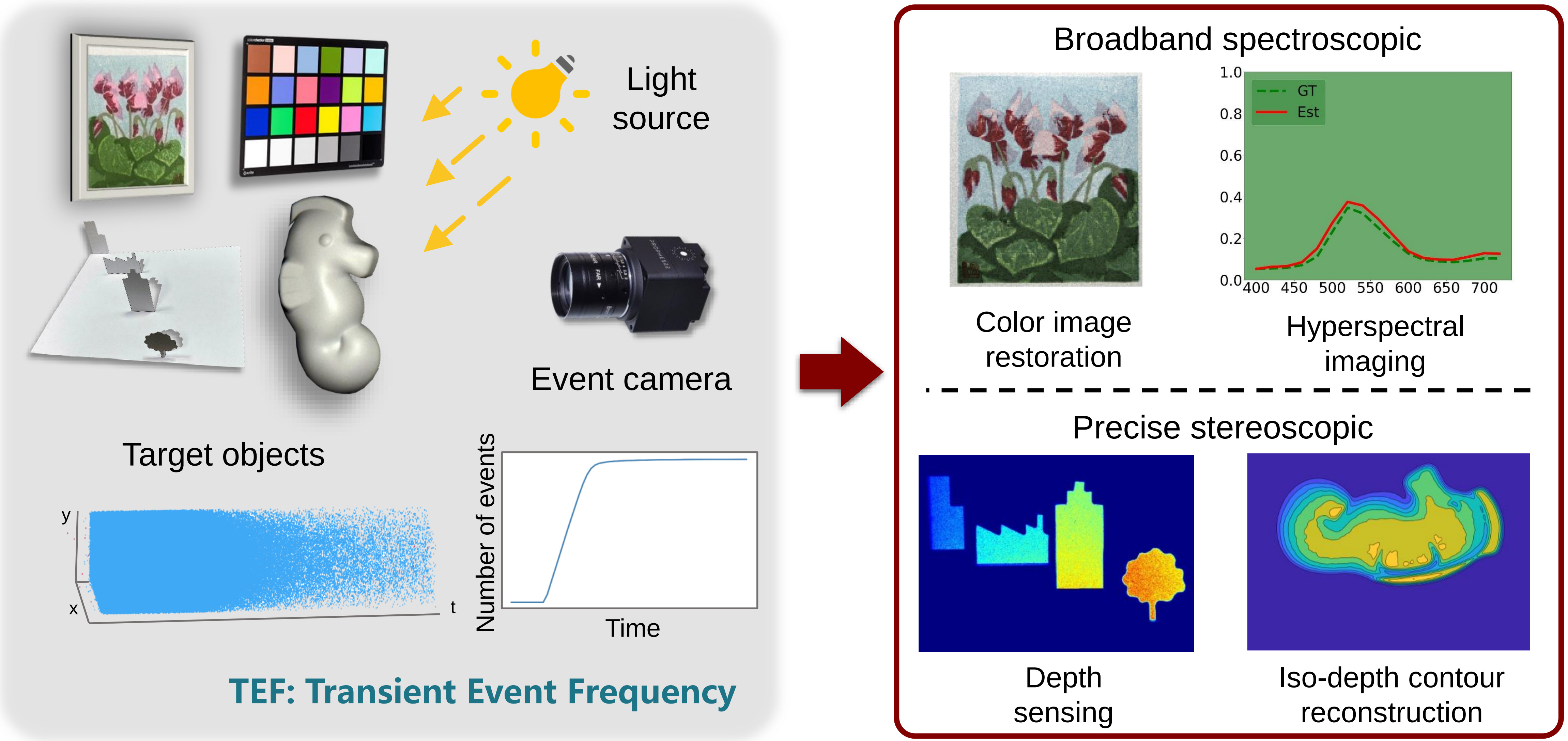


Hyperspectral imaging

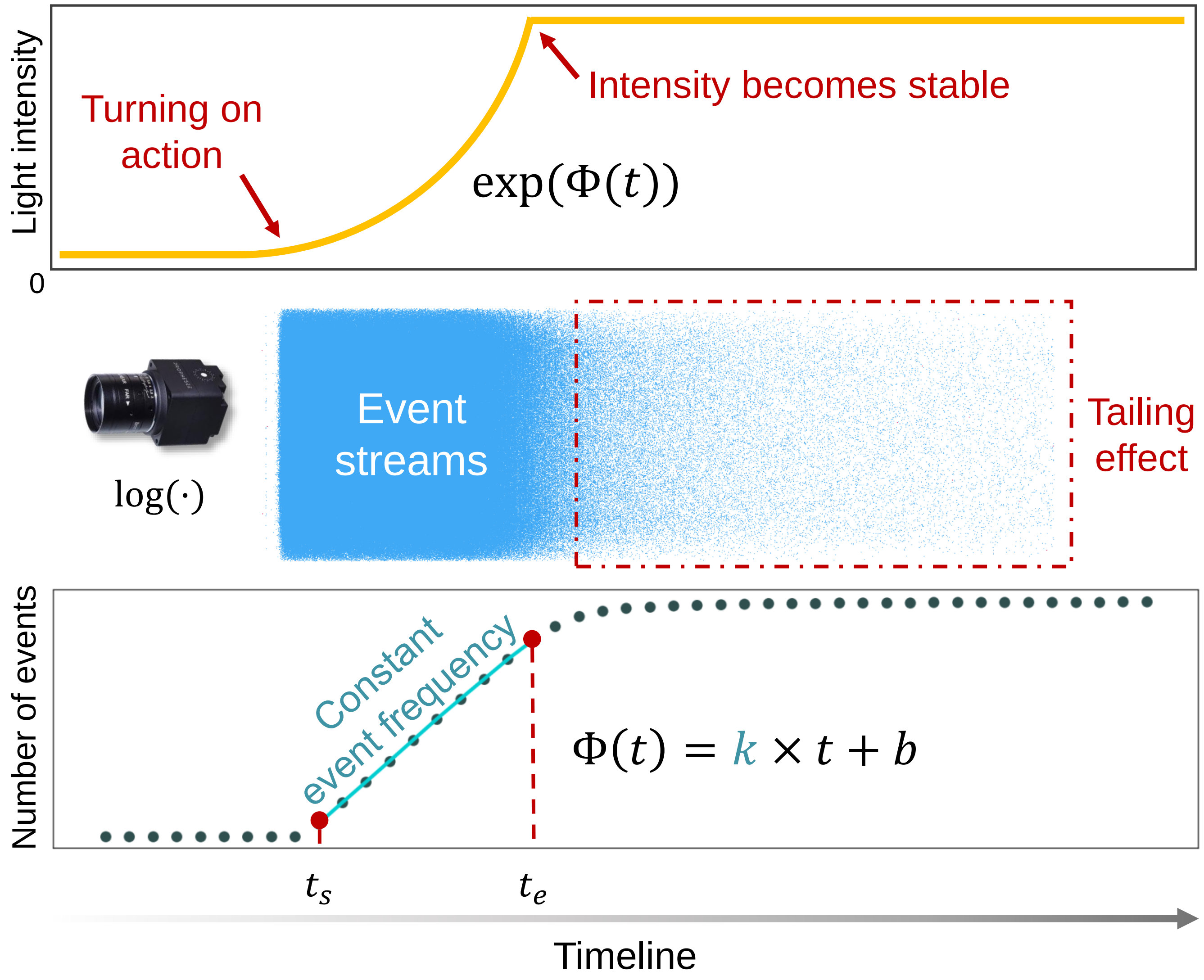


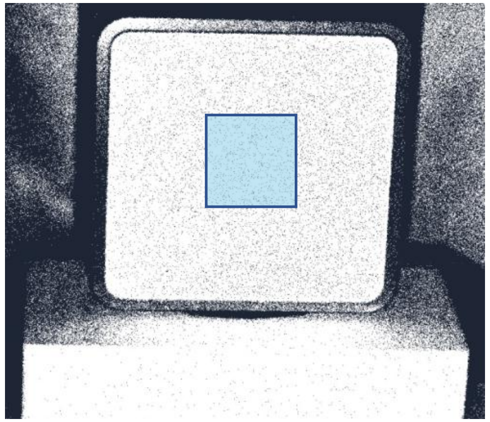
Iso-depth contour reconstruction

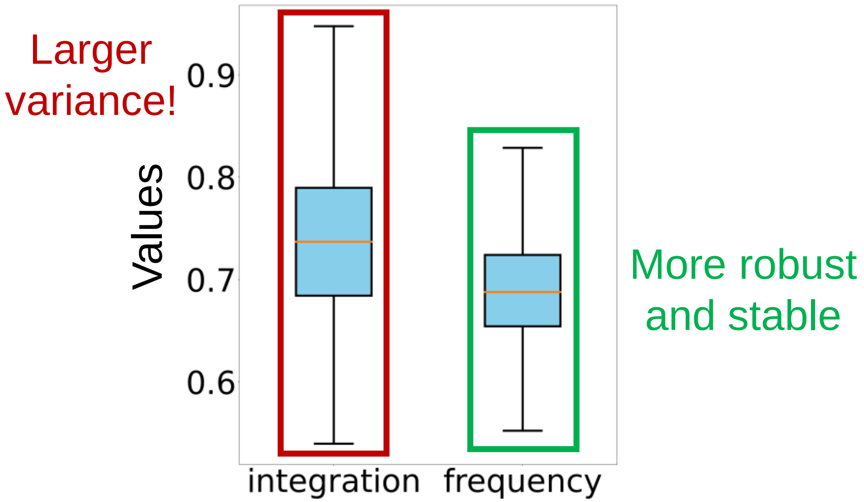
Overview of TEF (Transient Event Frequency)



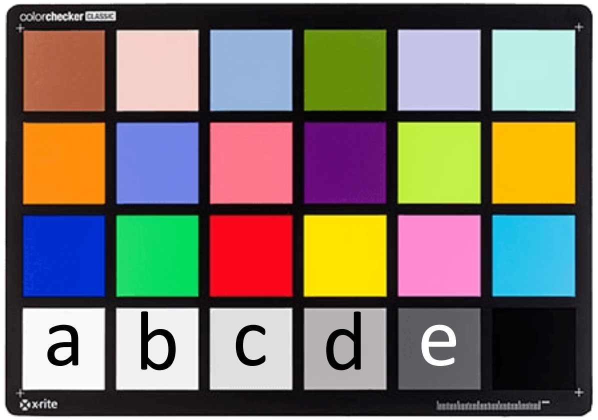
TEF: Transient Event Frequency



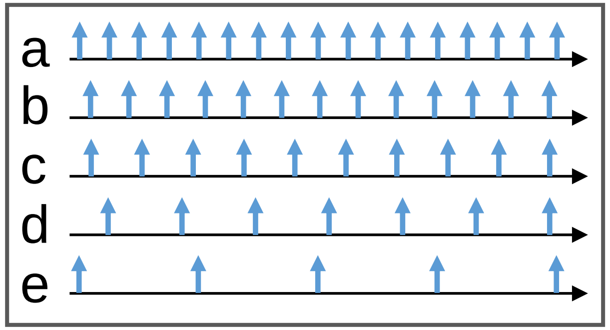




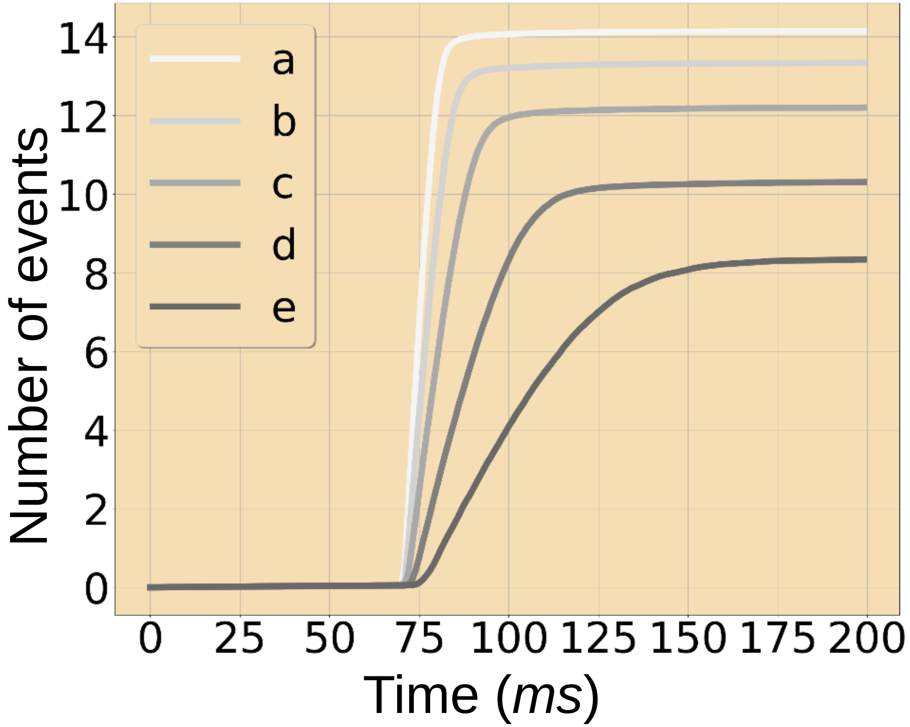
TEF: Transient Event Frequency



ColorChecker



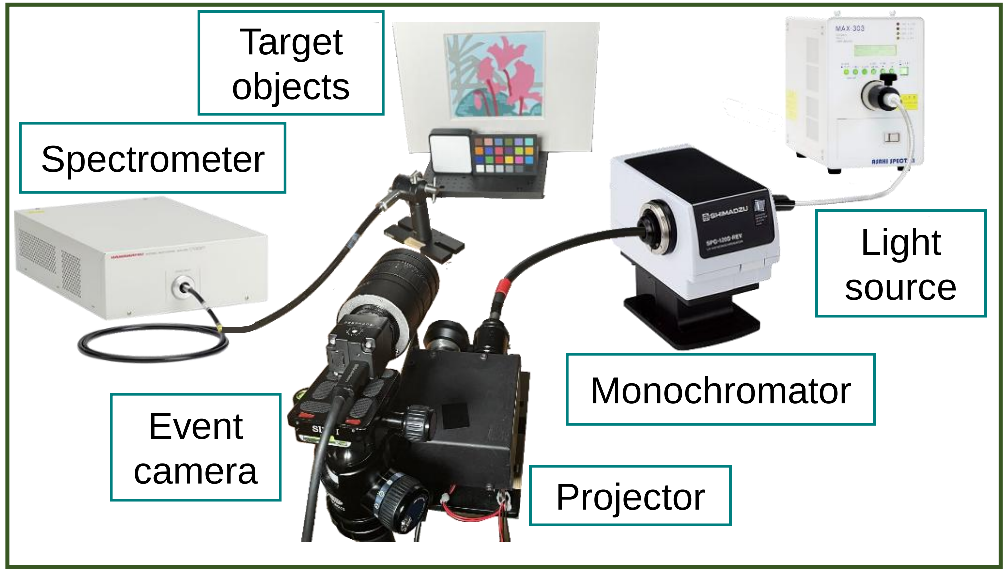
Positive events triggered in different patches



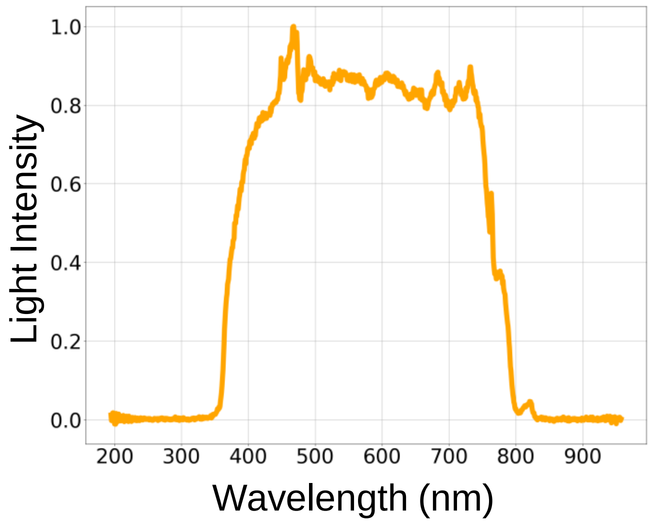
$$\Phi(t) = k \times t + b$$

Larger slope factor k represents higher radiance value.

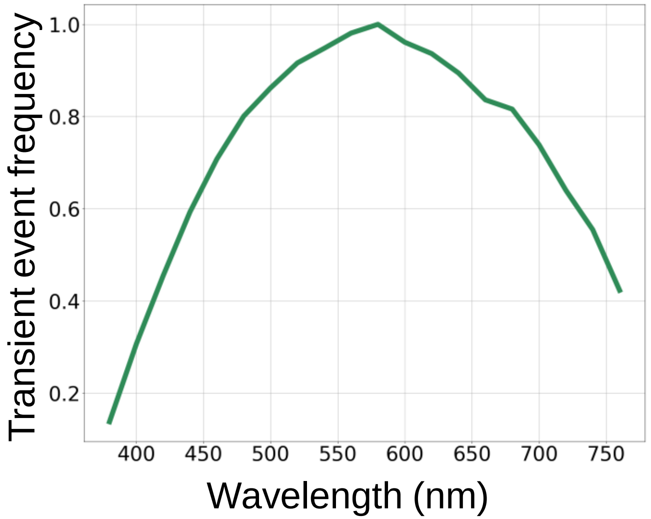
Spectral Response Function of Event Camera



Experimental setup



Spectral distribution of light source



Spectral response curve of Prophesee EVK4 (used for the following spectroscopic experiments)

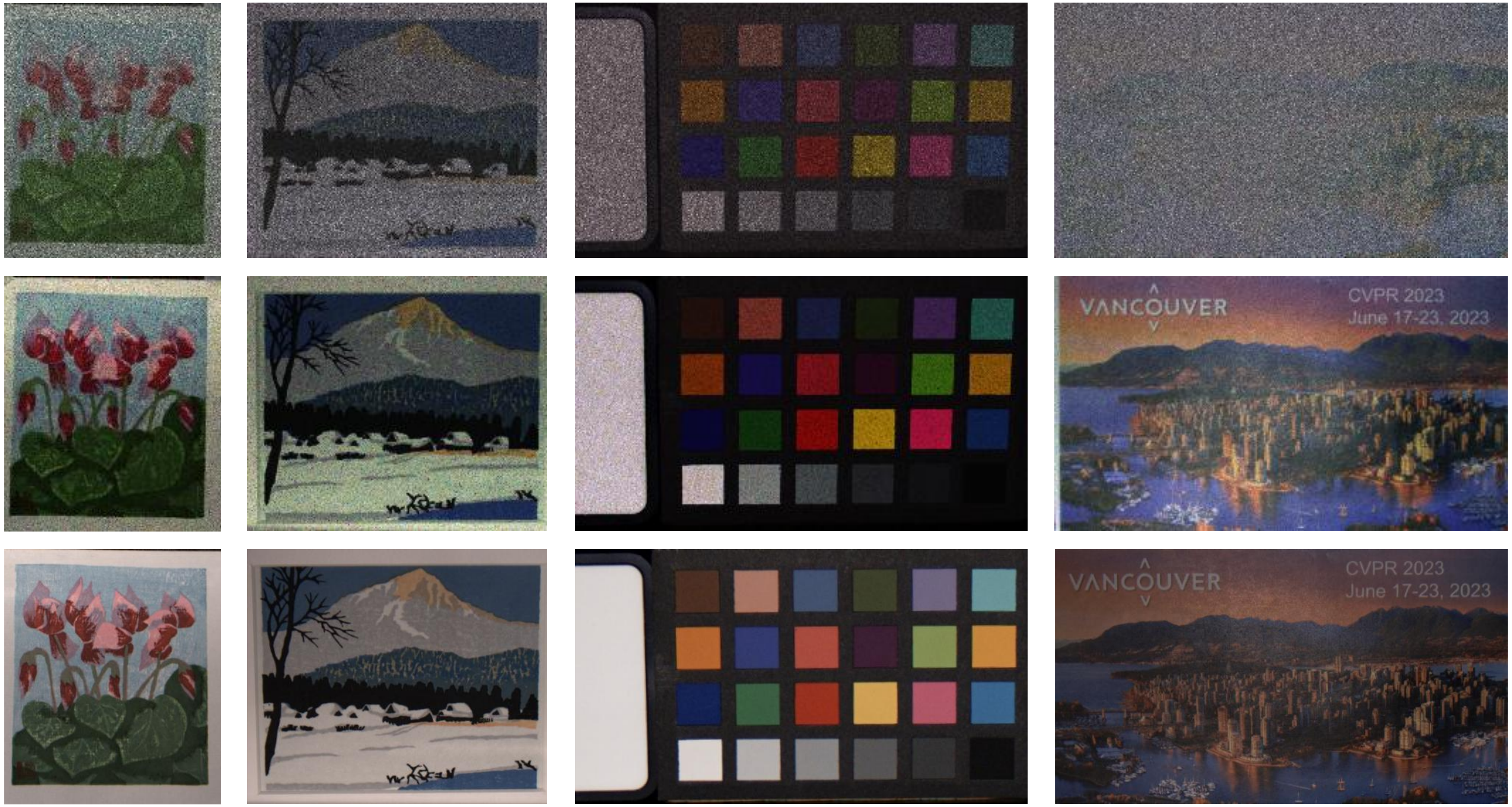
Experiments

- Color image restoration

Integration-based results (baseline)

TEF-based results (ours)

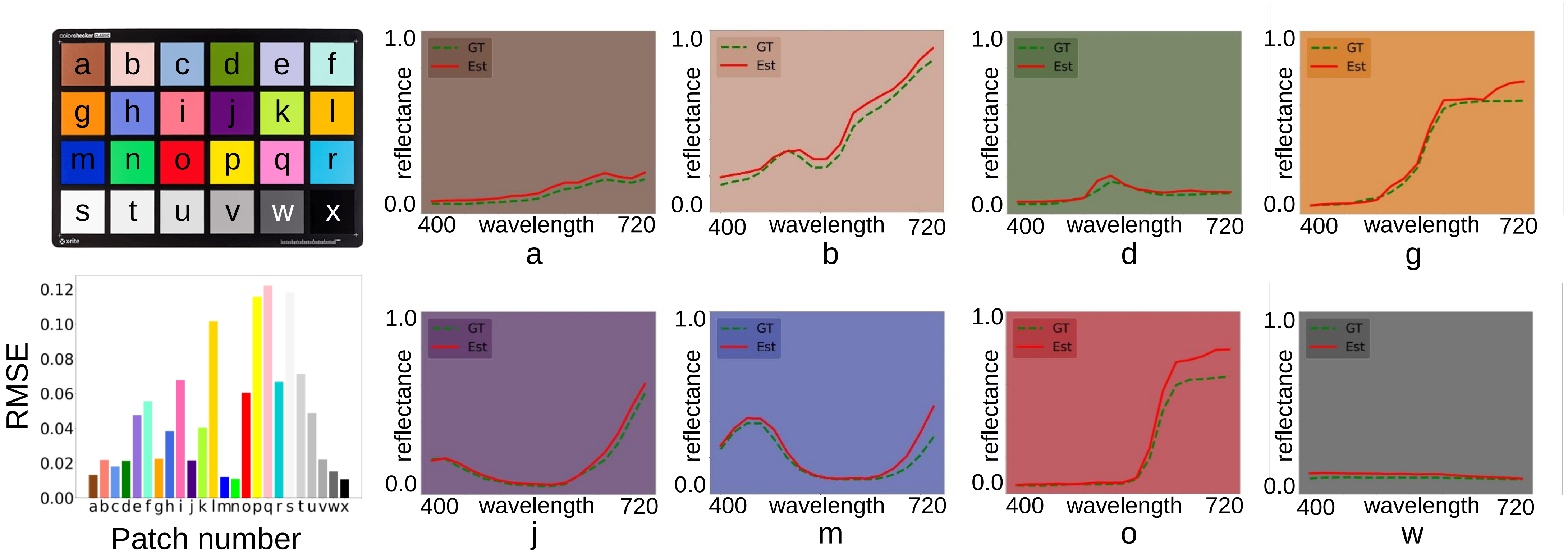
RGB camera (reference)



Postcard | Painting | ColorChecker | Printing paper

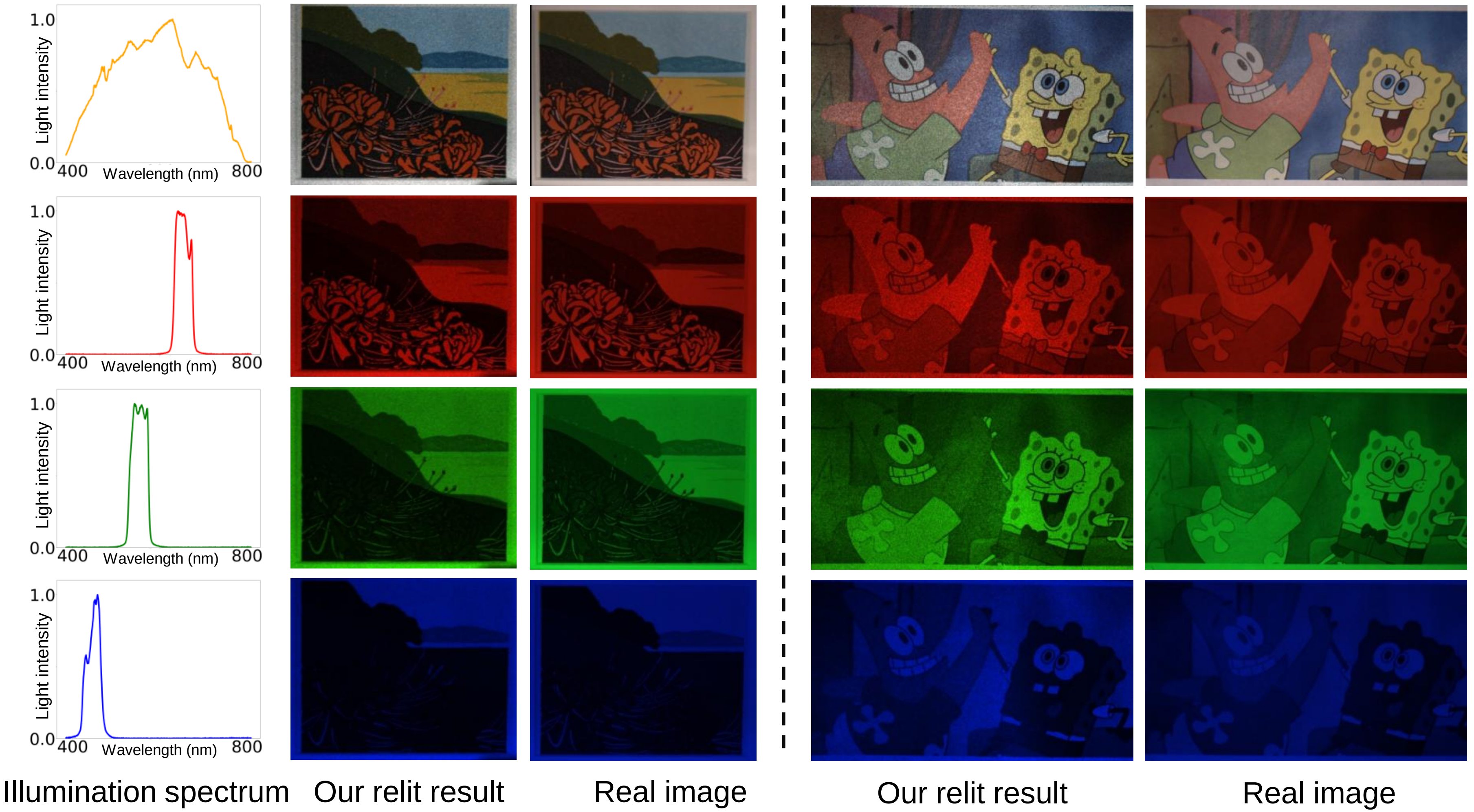
Experiments

- Hyperspectral imaging



Experiments

- Image relighting

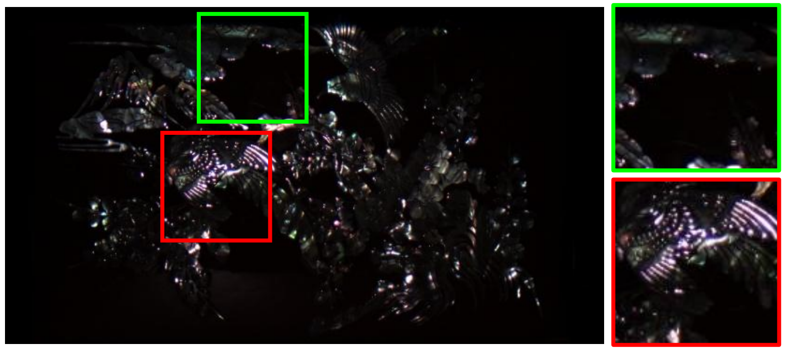


Experiments

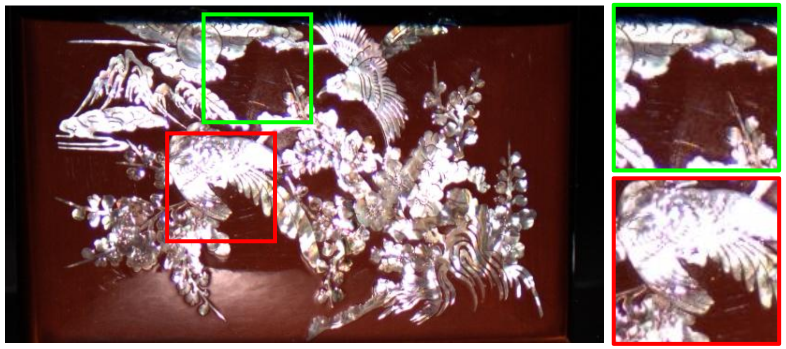
- High dynamic range example

A box surface with specular texture

Conventional camera

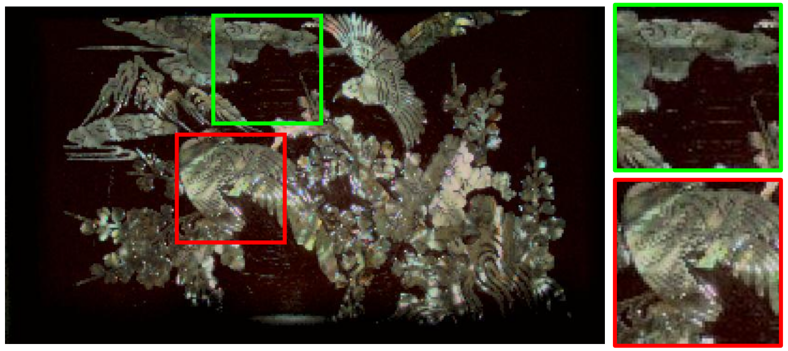


Short exposure



Long exposure

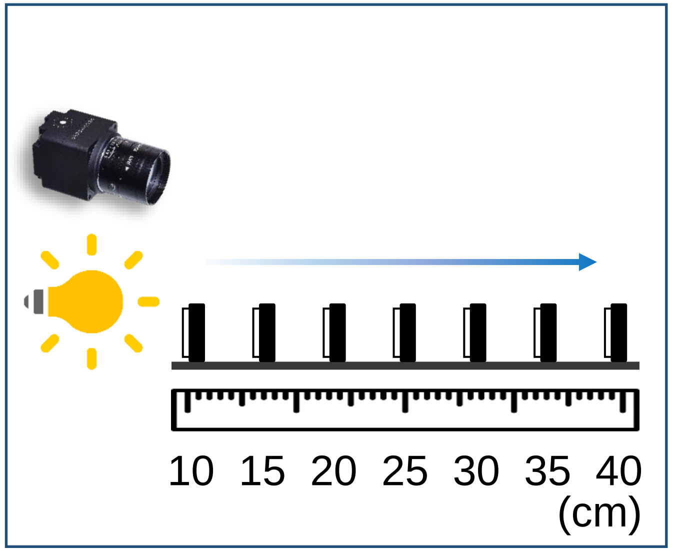
Event camera



Our relit result

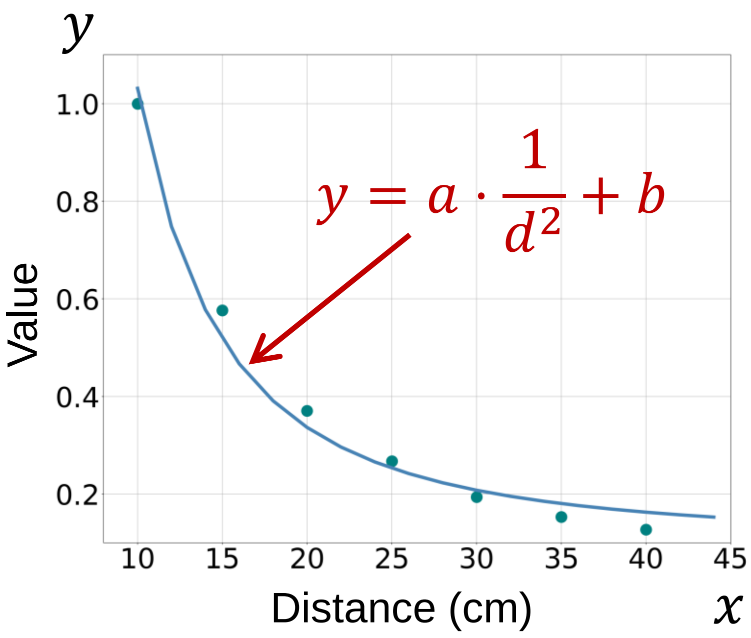
Experiments

- Depth sensing

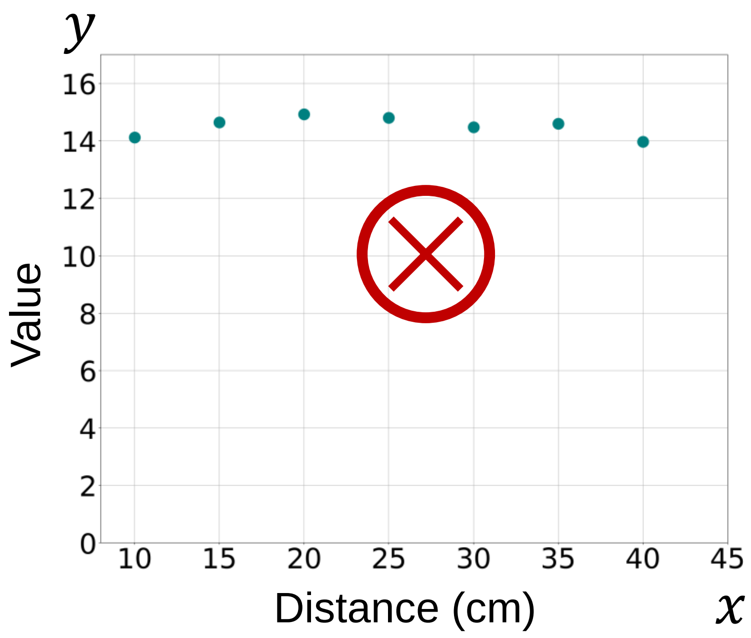


Experimental setup

$$\boxed{L_d = L \cdot \frac{1}{4\pi d^2}}$$



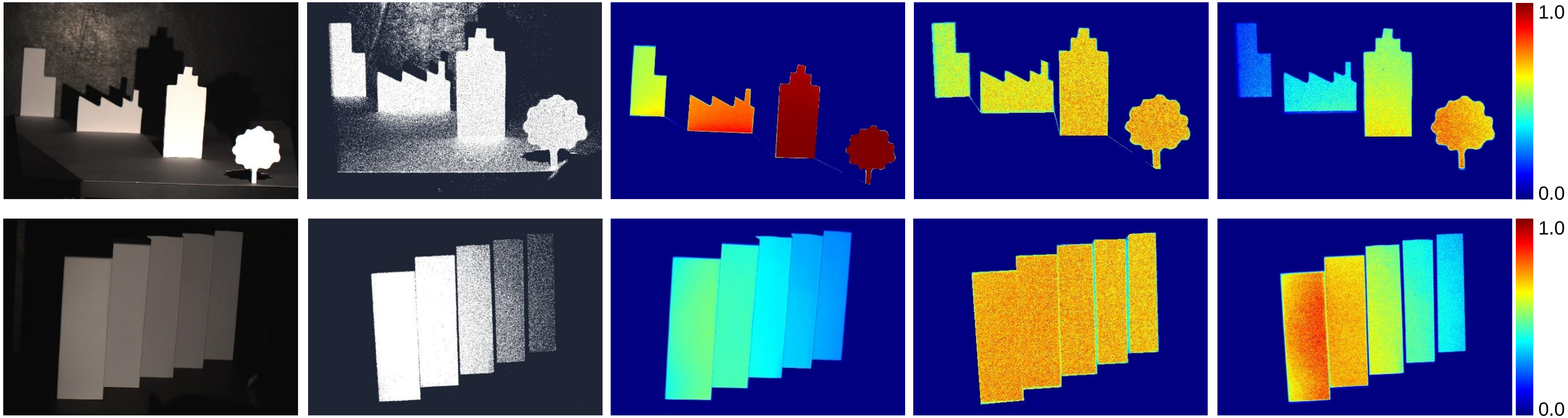
TEF-based recordings



Integration-based recordings

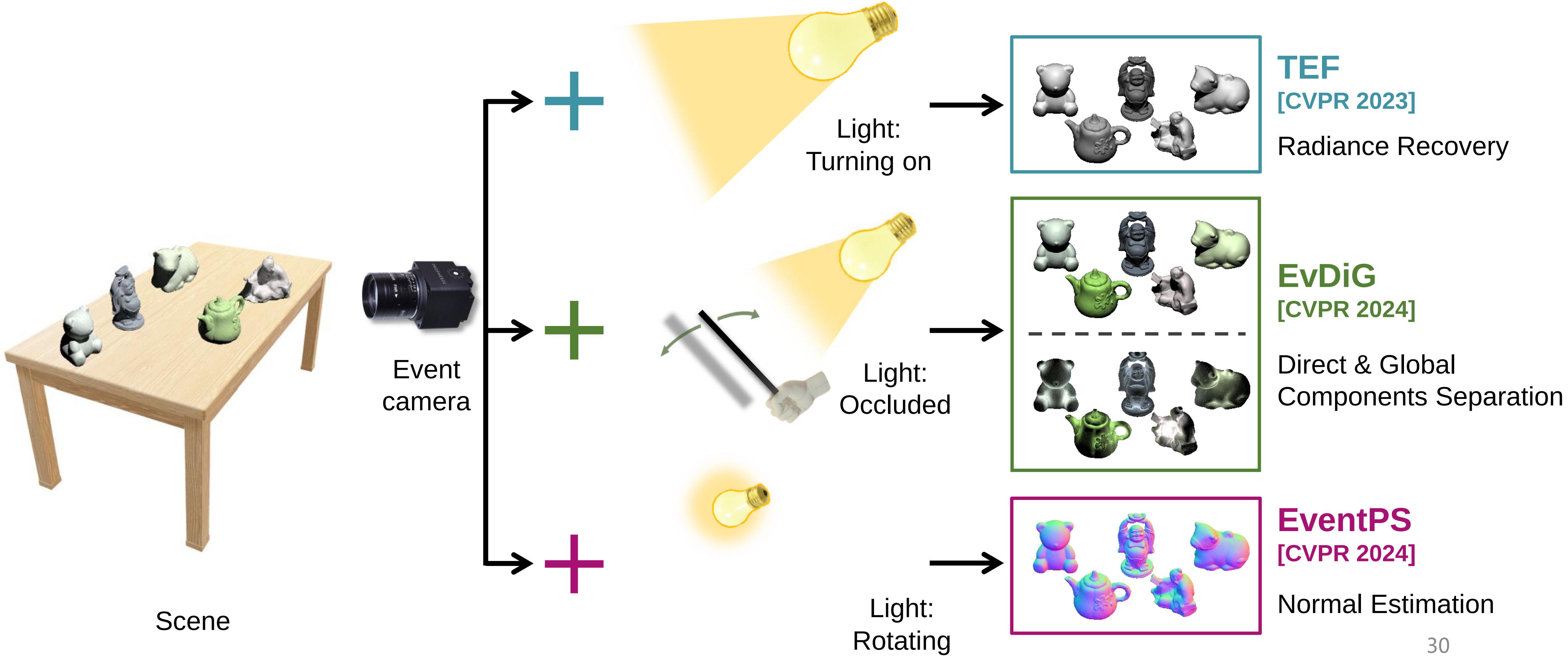
Experiments

- Depth sensing

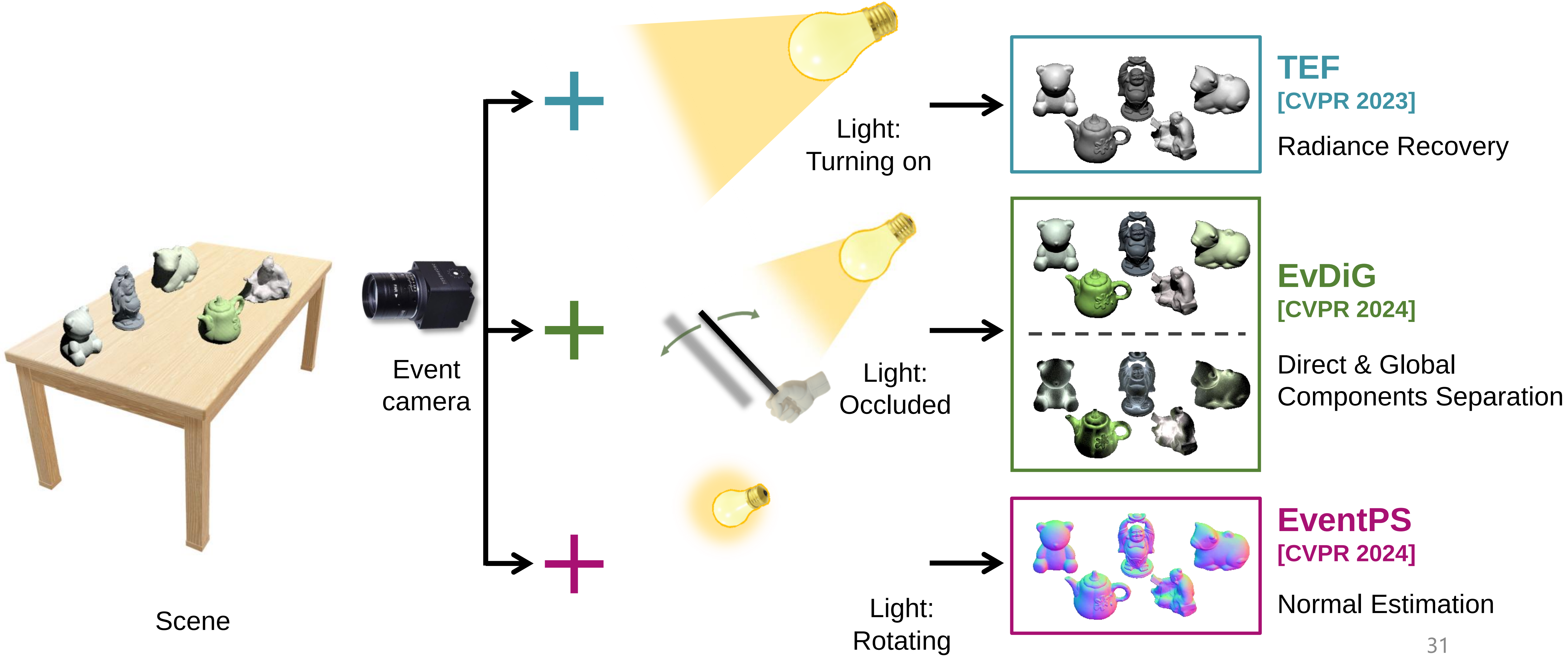


Target object | Event signals | Conventional camera result | Integration (baseline) | TEF (ours)

Event Photometry



Event Photometry

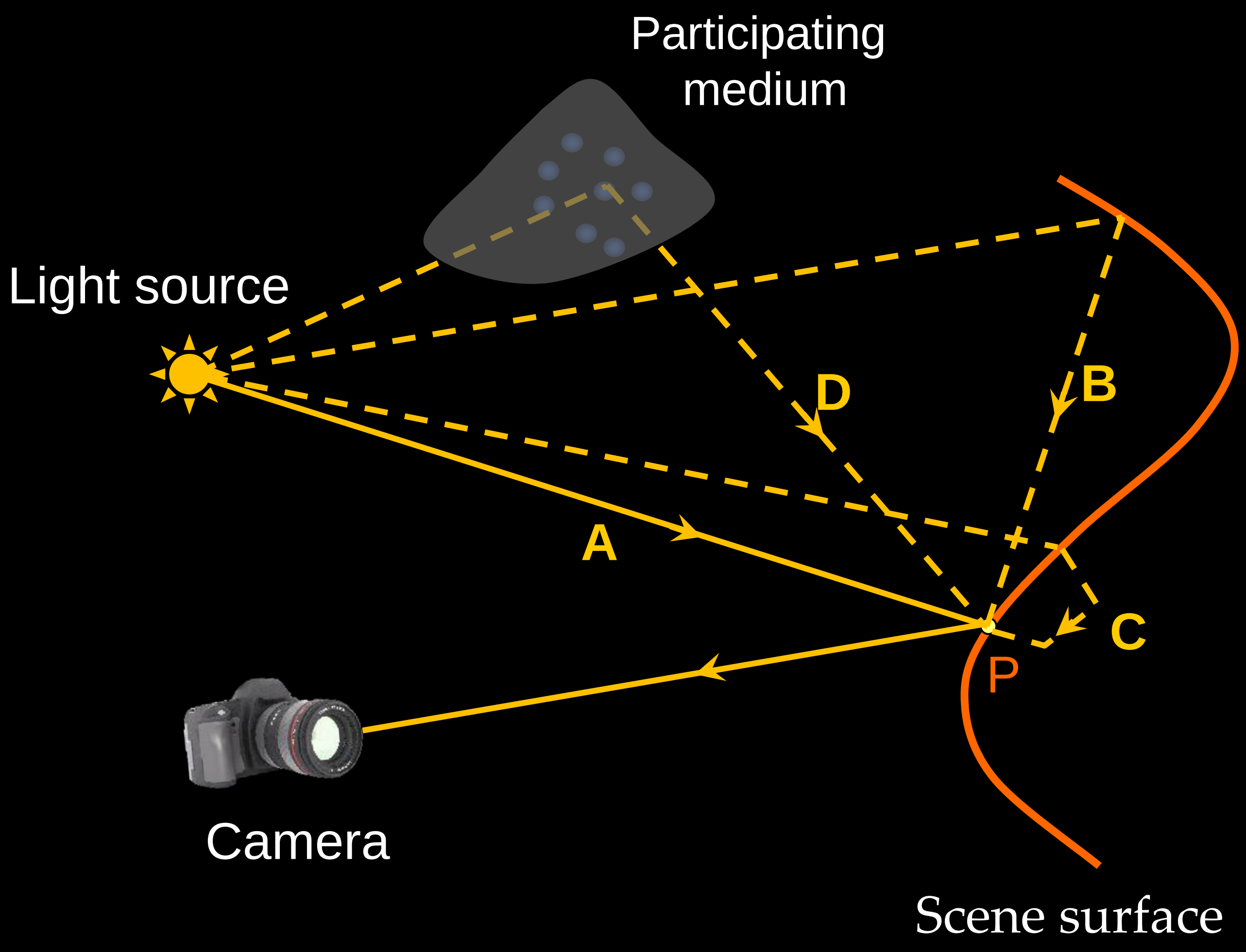


EvDiG: Event-guided Direct and Global Components Separation

Xinyu Zhou   Peiqi Duan   Boyu Li   Chu Zhou   Chao Xu   Boxin Shi

Oral / Full Scores in Final Review

Direct and Global Illumination



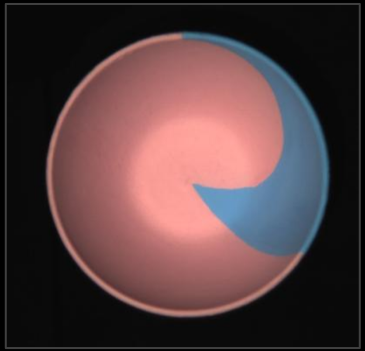
A : Direct

B : Interreflection
C : Subsurface
D : Volumetric

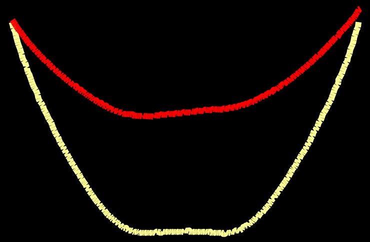
*Slide courtesy of Shree Nayar

• If Global Illumination is Ignored …

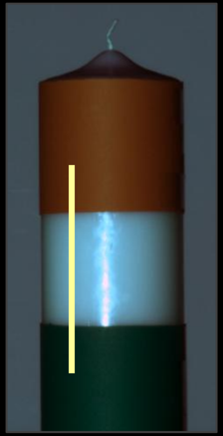
Photometric stereo



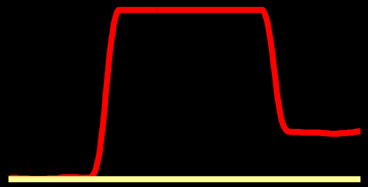
[Nayar et al. 91, 06]



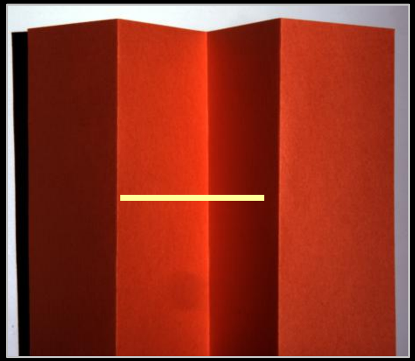
Shape from illumination defocus



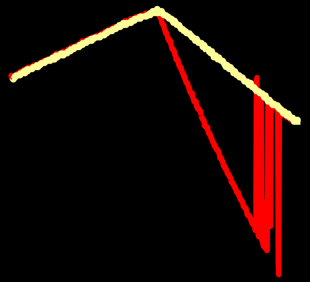
[Gupta et al. 09]



Structured light triangulation



[Gupta et al. 11]

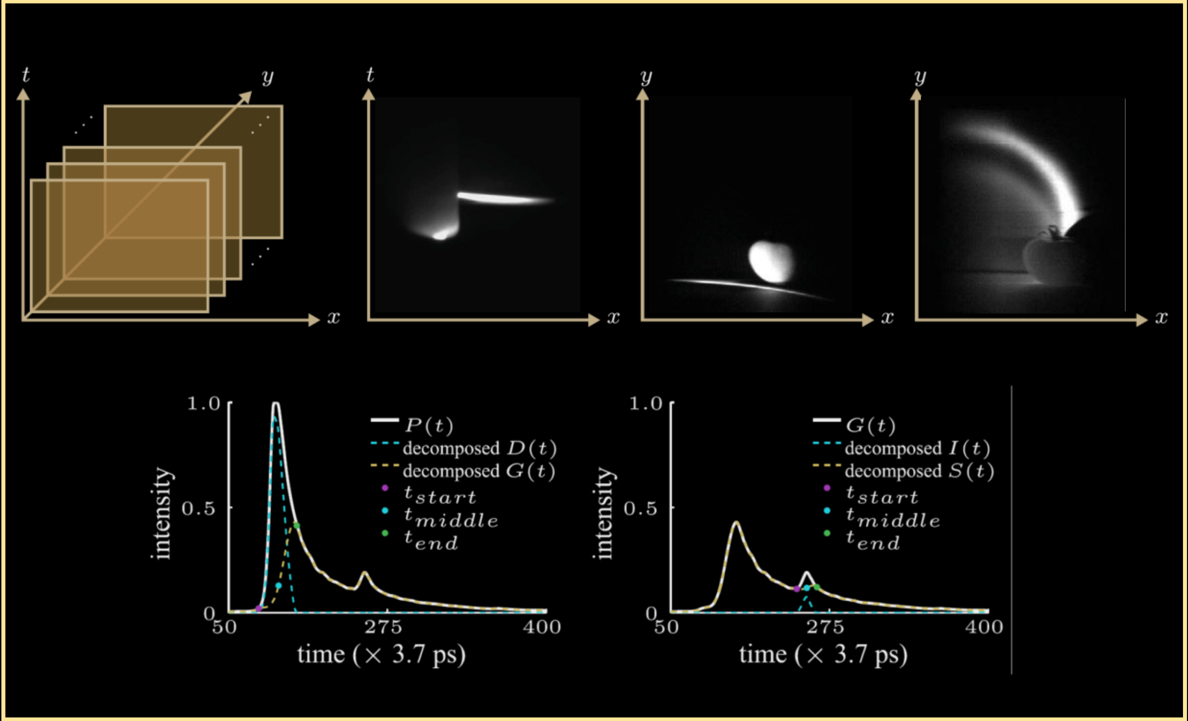


Estimated Shape

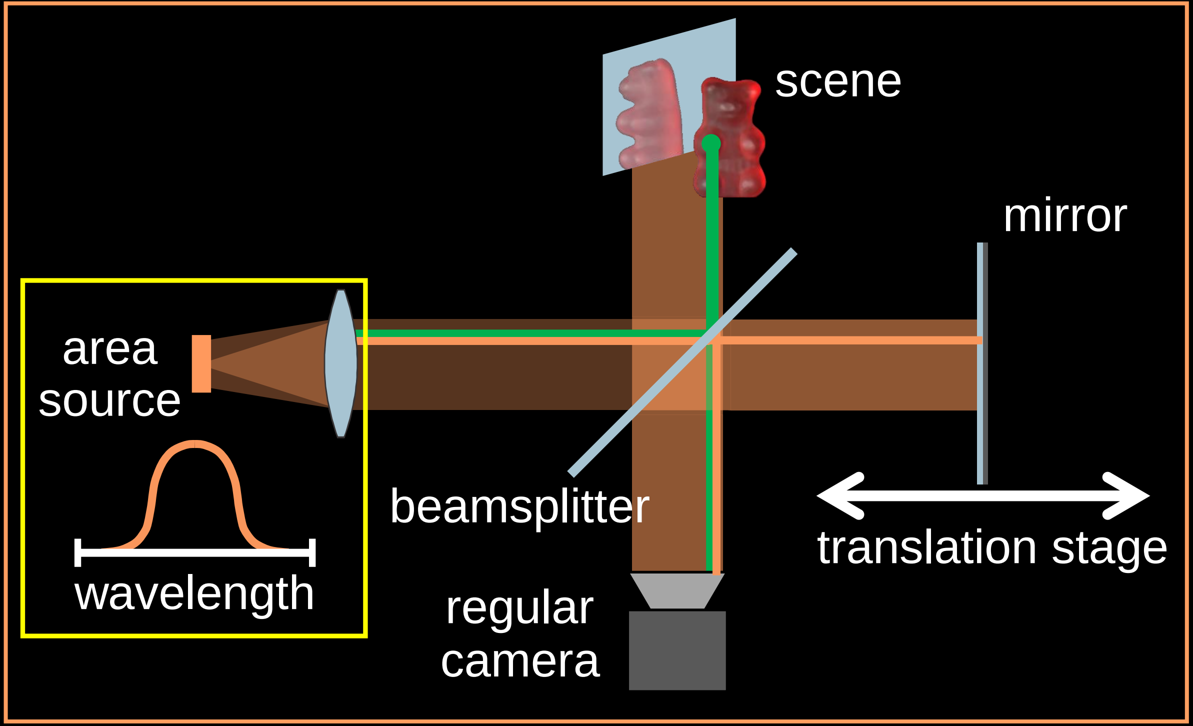
Ground Truth

*Slide courtesy of Mohit Gupta

● Current Separation Methods



Time of flight [Wu et al. 12]



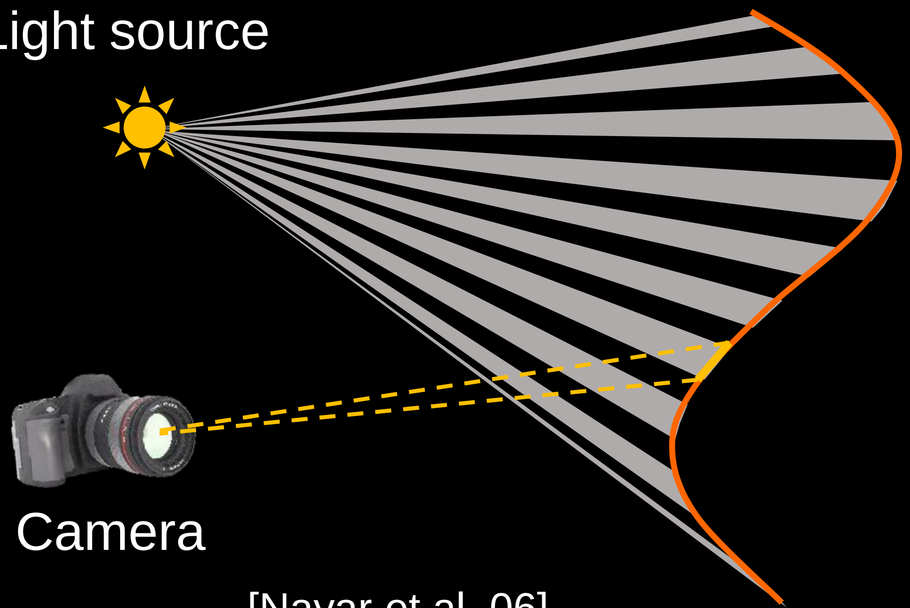
Optical interferometry [Gkioulekas et al. 15]

Time-resolved methods

Distinguish the different arrival time

High cost of special devices and long data capture time

Current Separation Methods



[Nayar et al. 06]

fraction of activated source elements

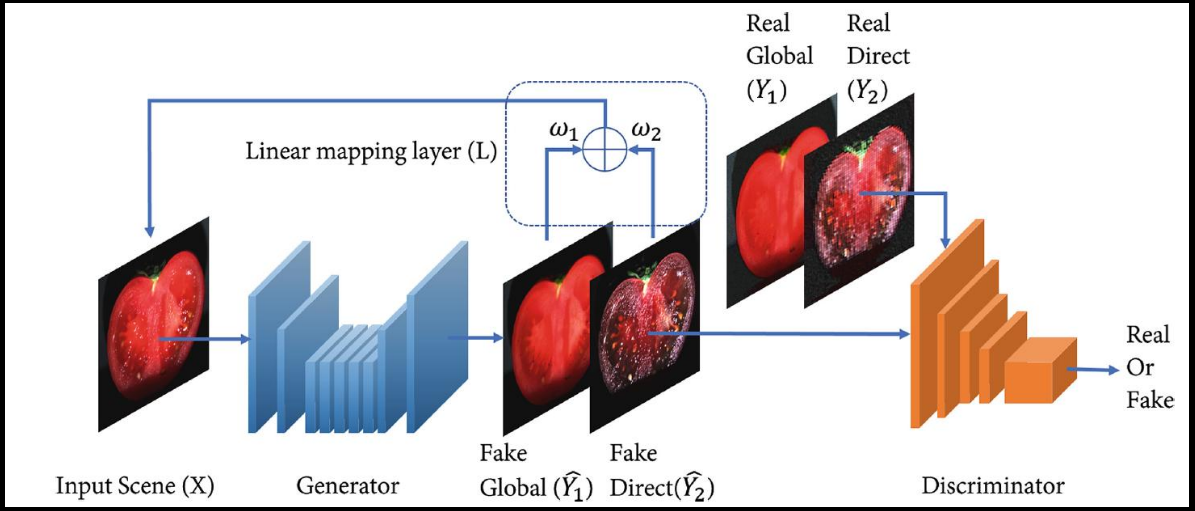
$$\boxed{L_{+}[i] = L_{d}[i] + \alpha L_{g}[i]}$$

$$\boxed{L_{-}[i] = (1-\alpha) L_{g}[i]}$$

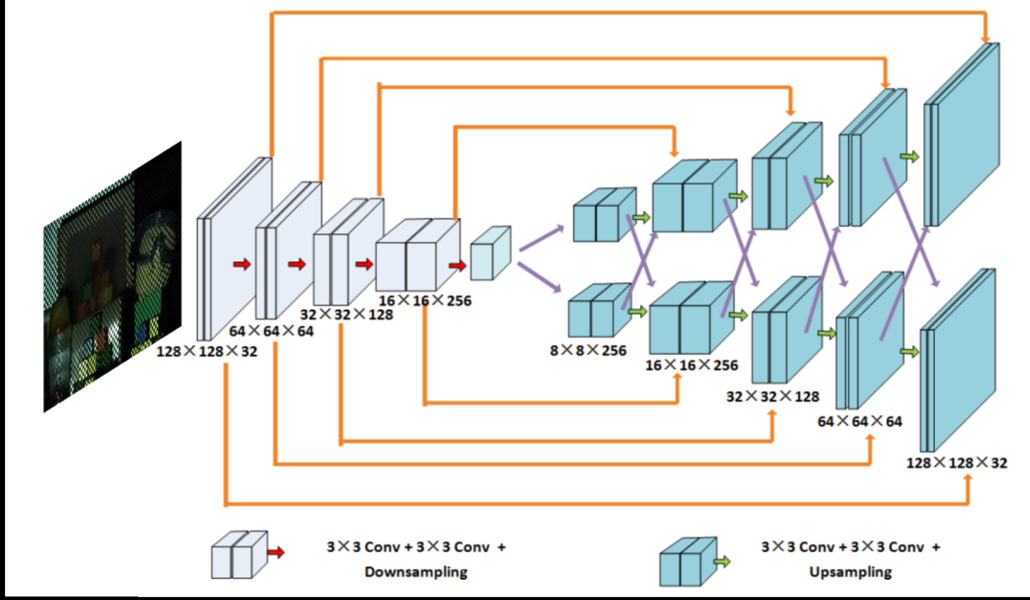
High frequency illumination pattern

Requires multiple images

● Current Separation Methods



Single scene image [Nie et al. 18]



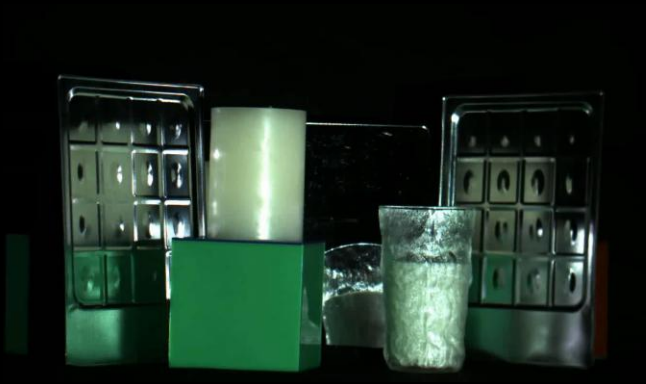
Single pattern image [Duan et al. 20]

Single-image methods

Low separation quality due to fewer physical cues

Motivation

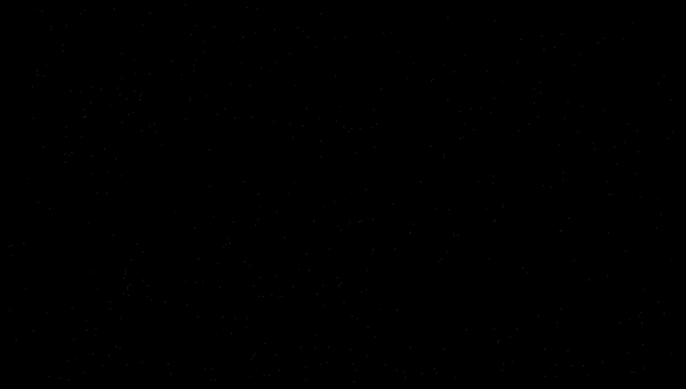
Low temporal resolution



RGB frame

30× slower

High temporal resolution



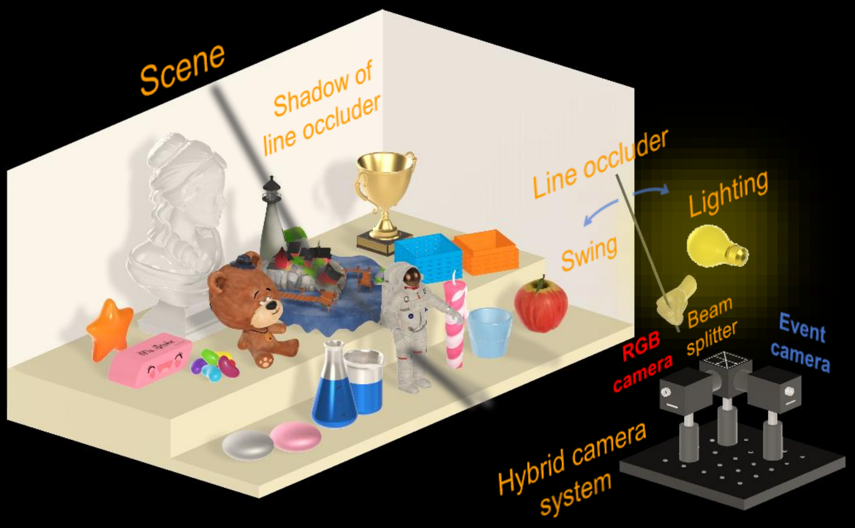
Event

Event-guided separation

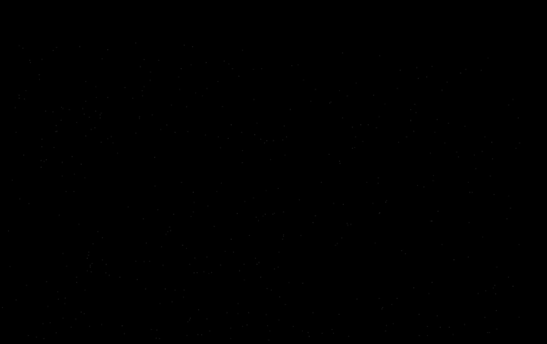
High speed events measure the dynamic intensity changes induced by moving shadows over time, **reducing** data capture **time**

EvDiG Settings

To **dig** out Event cameras’ unique properties for Direct and Global components separation



Single scene image



Event

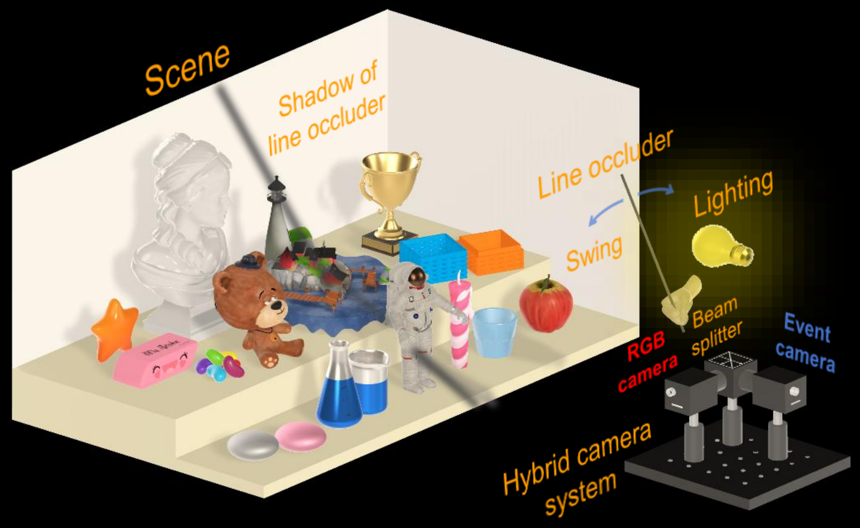
Fast, HDR, and sparse

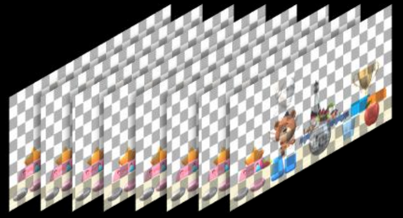
Direct component

Global component

● EvDiG Settings

To **dig** out **Ev**ent cameras' unique properties for **Di**rect and **G**lobal components separation



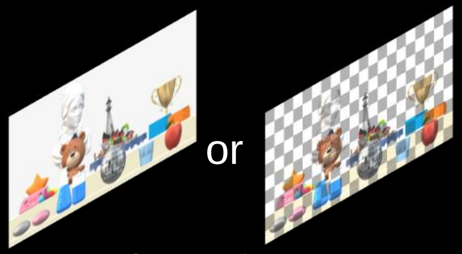


Multi-frames with patterns

25 images / 833ms

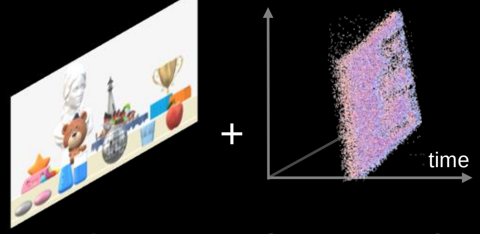
Multi-frames with shadow

> 30 images / 1000ms



Single-frame

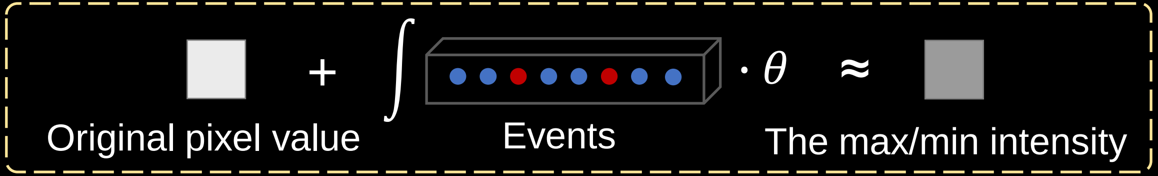
1 image / 33ms

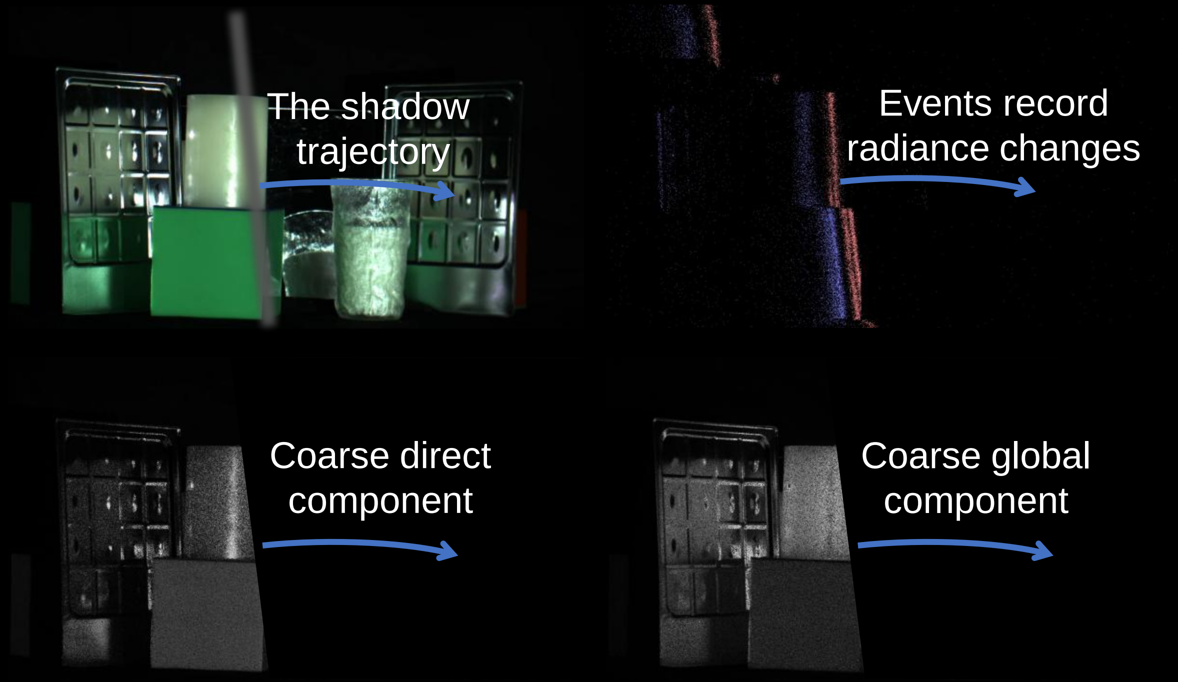


Single-frame and events (Ours)

1 image / 33ms

Formulation





- Separation through **event accumulation**
  - Given the grayscale scene image at t_0 and the events, reconstruct the latent image at any time during the shadow sweep

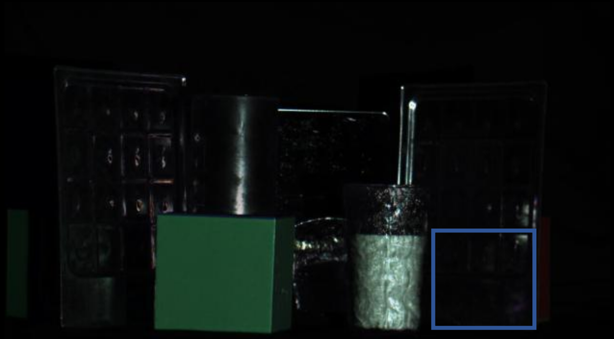
$$\log(\mathbf{I}^t) = \log(\mathbf{I}^{t_0}) + \theta \cdot \sum \{e_k\}^{t_0 \to t}$$

  - Obtain maximum and minimum intensity in each pixel to get coarse separation results via

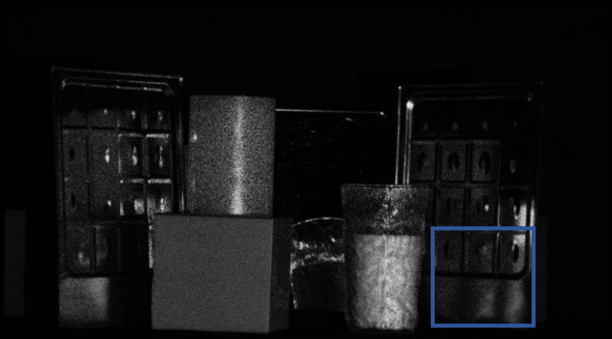
$$\mathbf{I}_{\mathrm{max}} = \mathbf{I}_{\mathrm{direct}} + \mathbf{I}_{\mathrm{global}}$$

$$\mathbf{I}_{\mathrm{min}} = \mathbf{I}_{\mathrm{global}}$$

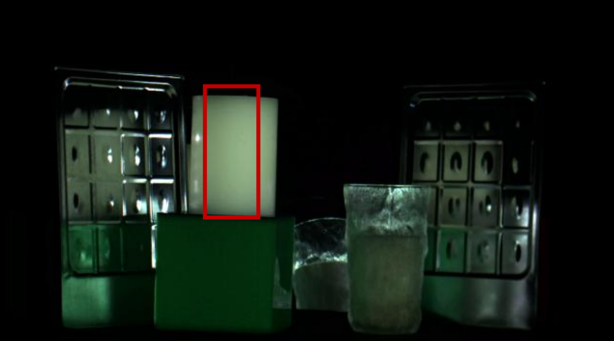
Problems in Coarse Results



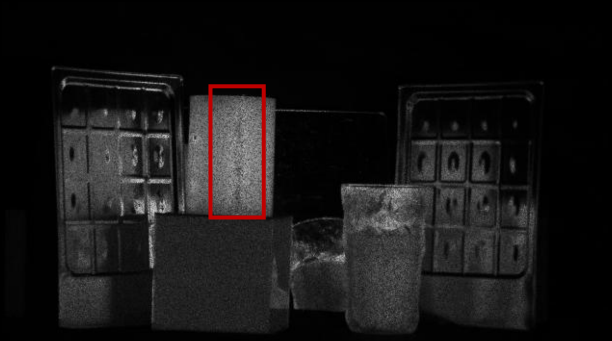
Reference direct



Coarse direct



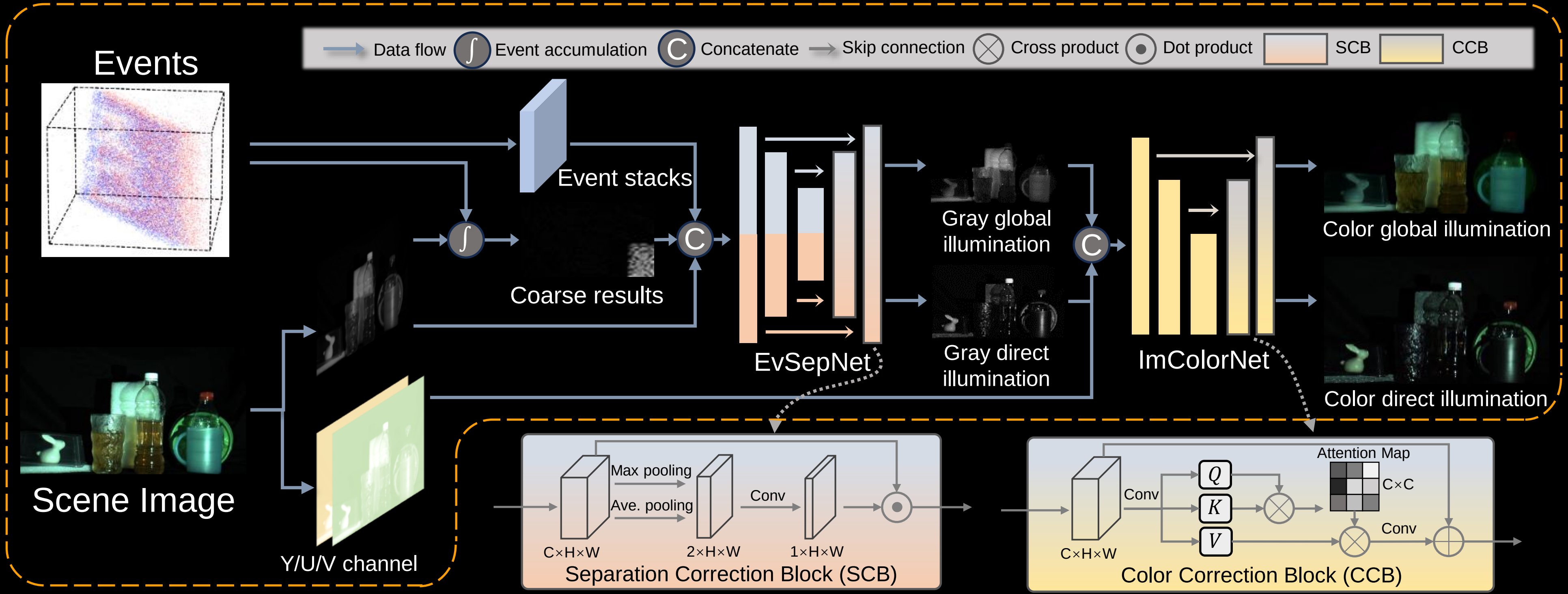
Reference global



Coarse global

- **Noisy**: Non-negligible sensor noise and spatial-temporal variant thresholds in event cameras
- **Inaccurate**: Quantization error in the event generation model
- **Monochromatic**: Event cameras record intensity changes information in grayscale

EvDiG Pipeline

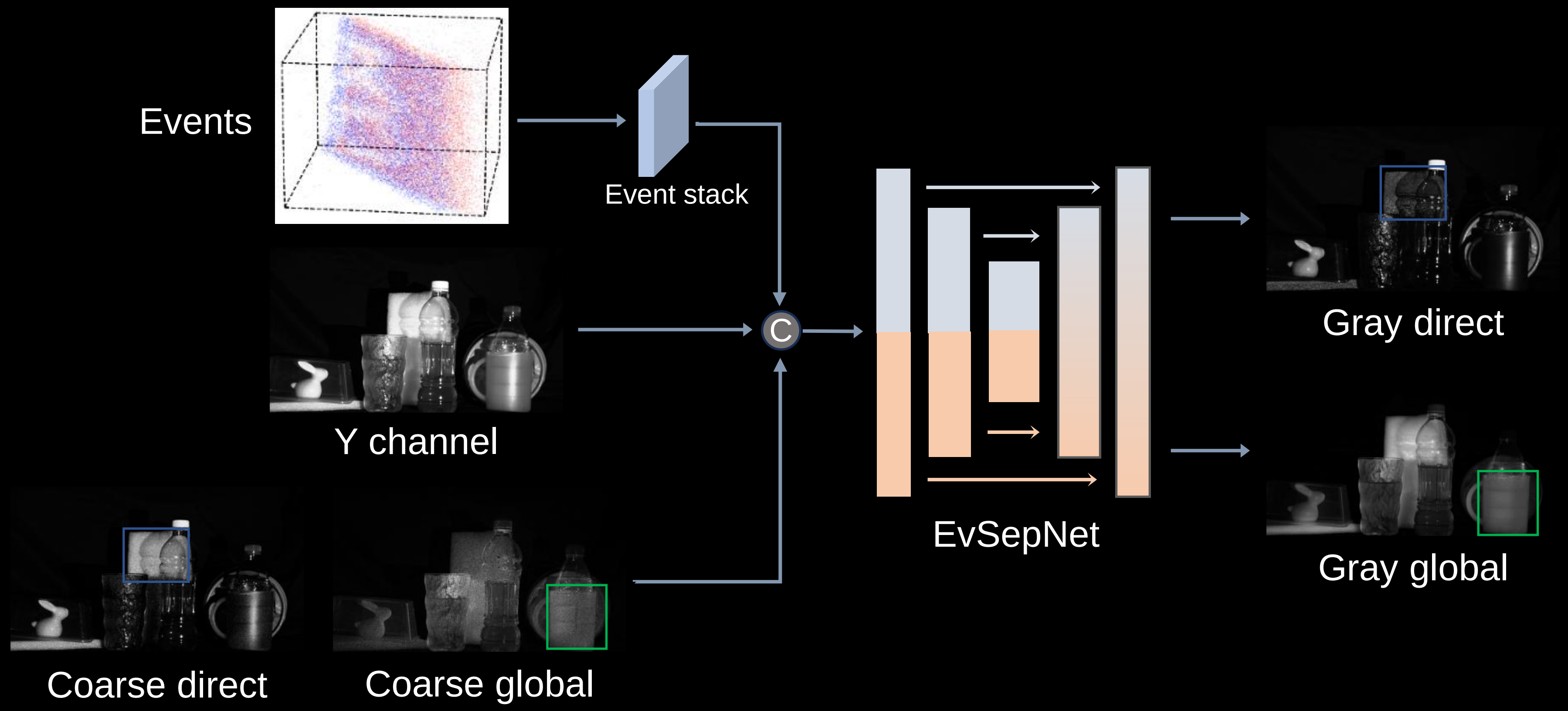


Coarse separation results → EvSepNet → ImColorNet

Refine coarse results

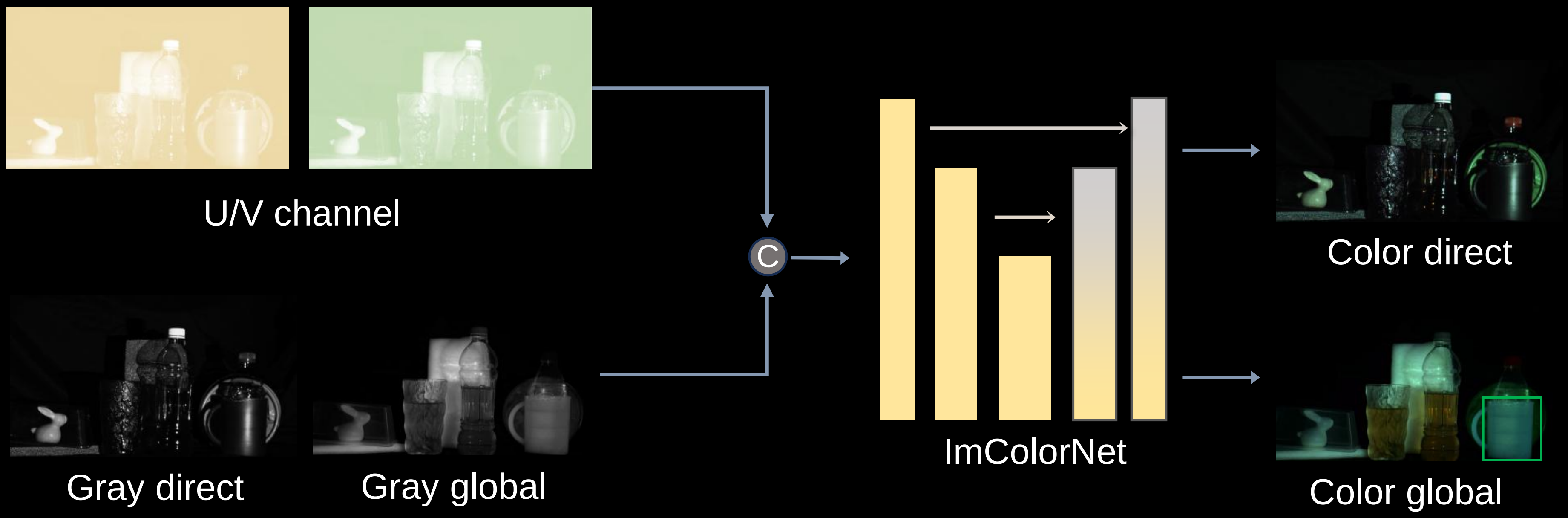
Chrominance compensation

EvDiG Pipeline



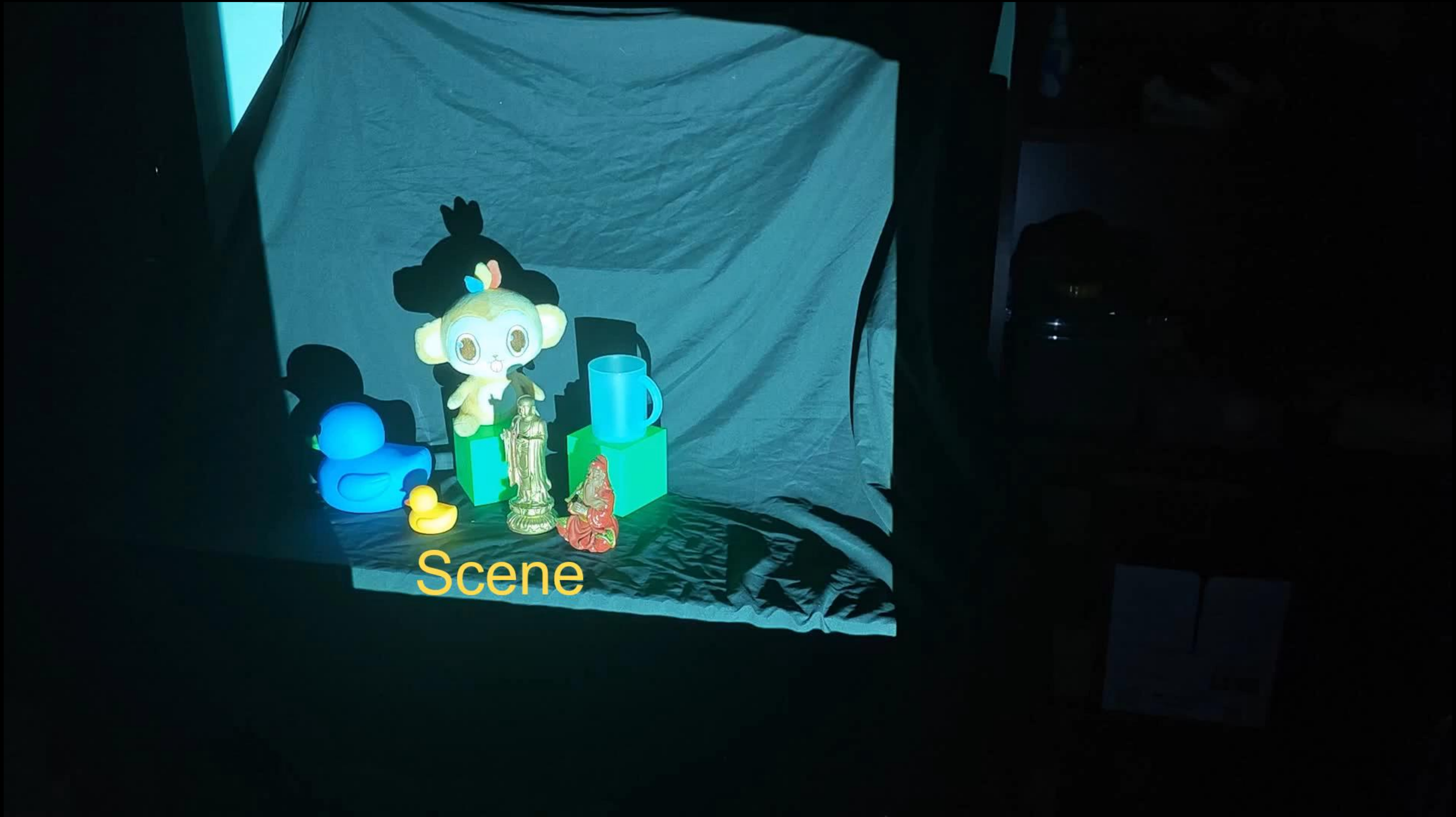
Remove noise and correct inaccurate separation regions

EvDiG Pipeline



Compensate chrominance information to get colorful and high-quality separation results

Data Capture

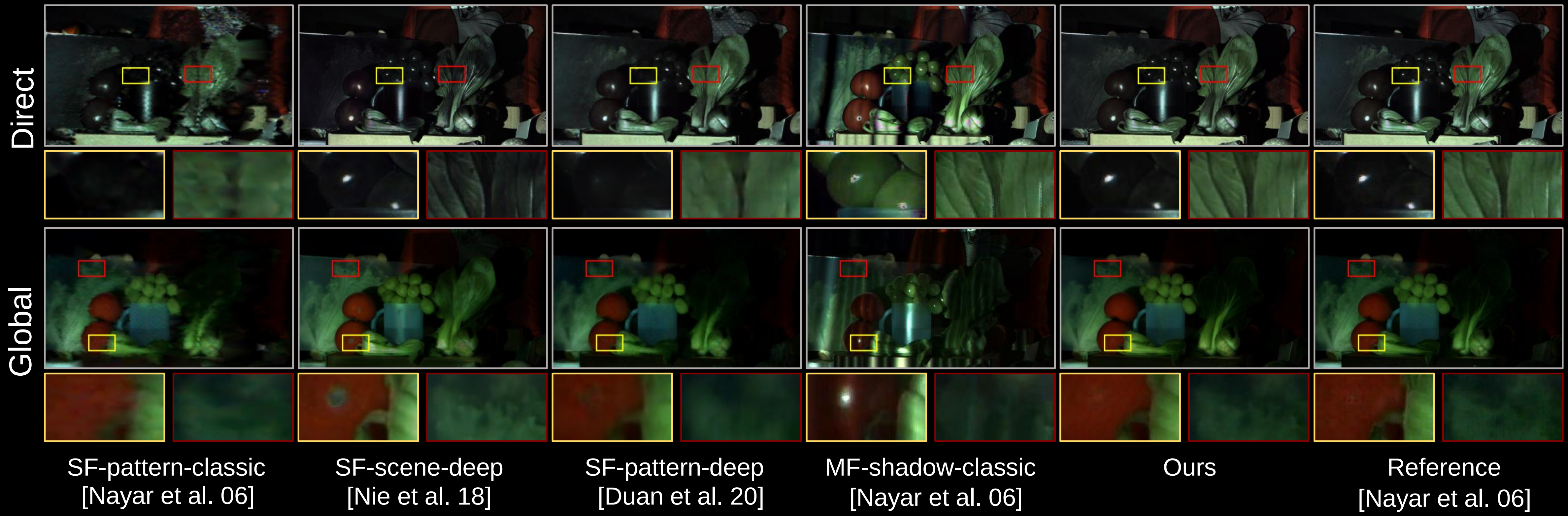


Results on Indoor Scenes

SF-pattern-classic [Nayar et al. 06] | SF-scene-deep [Nie et al. 18] | SF-pattern-deep [Duan et al. 20] | MF-shadow-classic [Nayar et al. 06] | Ours | Reference [Nayar et al. 06]

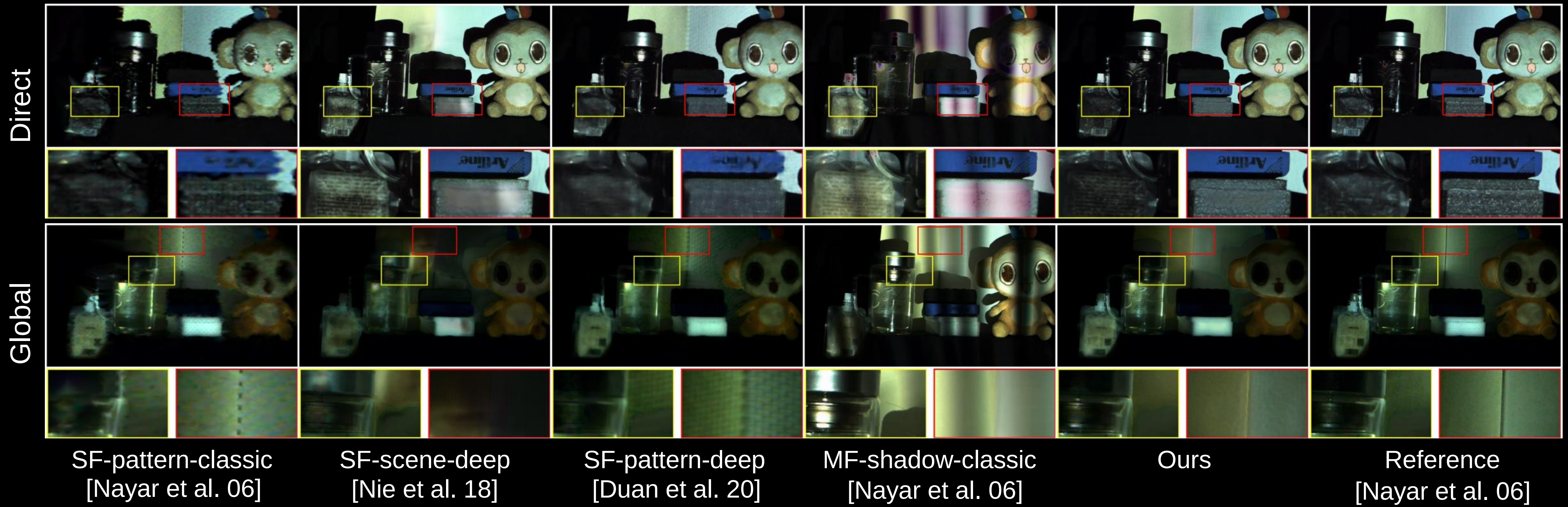
* SF (single-frame) / MF (multi-frame)

Results on Indoor Scenes



* SF (single-frame) / MF (multi-frame)

● Results on Indoor Scenes



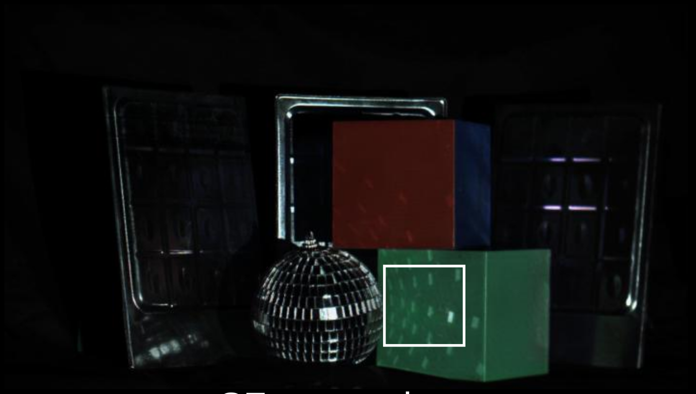
* SF (single-frame) / MF (multi-frame)

Results on Indoor Scenes

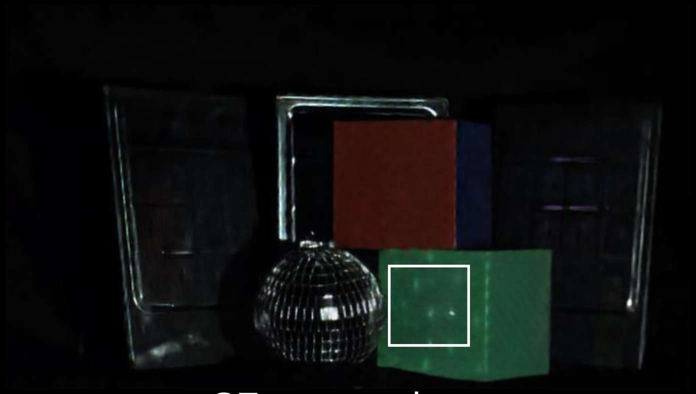
with high-frequency global illuminations

Direct

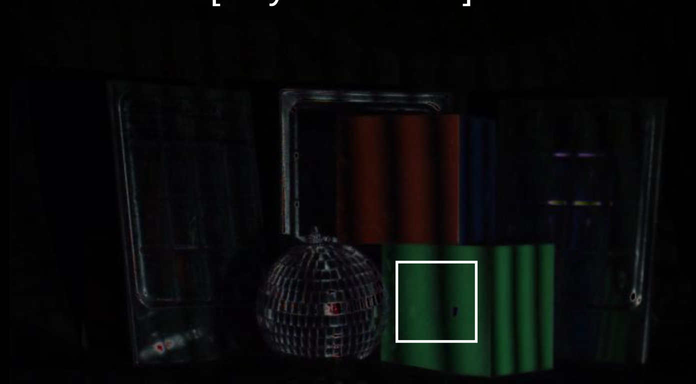
SF-pattern-classic
[Nayar et al. 06]



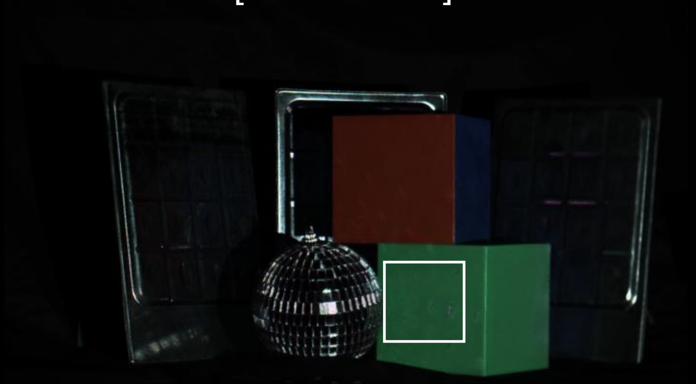
SF-scene-deep
[Nie et al. 18]



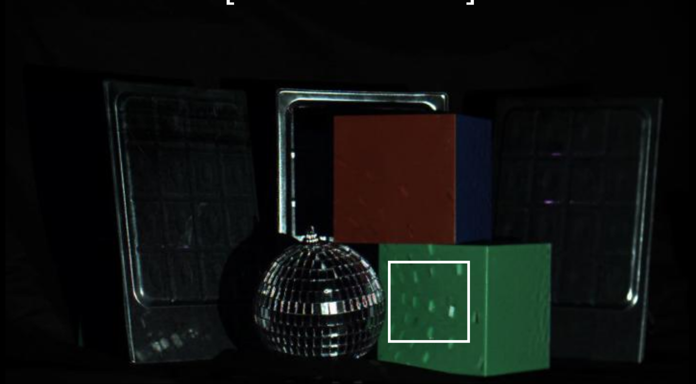
SF-pattern-deep
[Duan et al. 20]



MF-shadow-classic
[Nayar et al. 06]



Ours



Reference
[Nayar et al. 06]

• Results on Outdoor Scenes

Scene

Direct

Global

MF-shadow-classic [Nayar et al. 06]

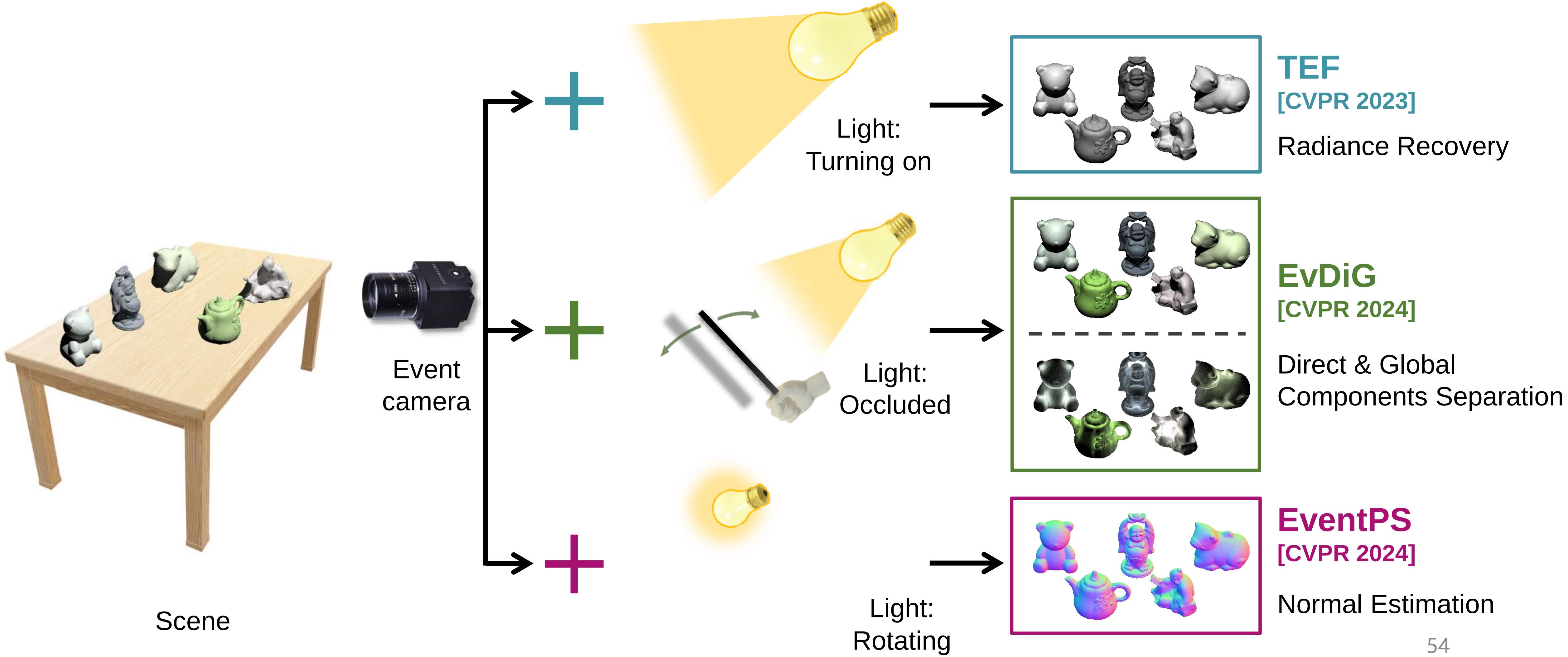
SF-scene-deep [Nie et al. 18]

Ours

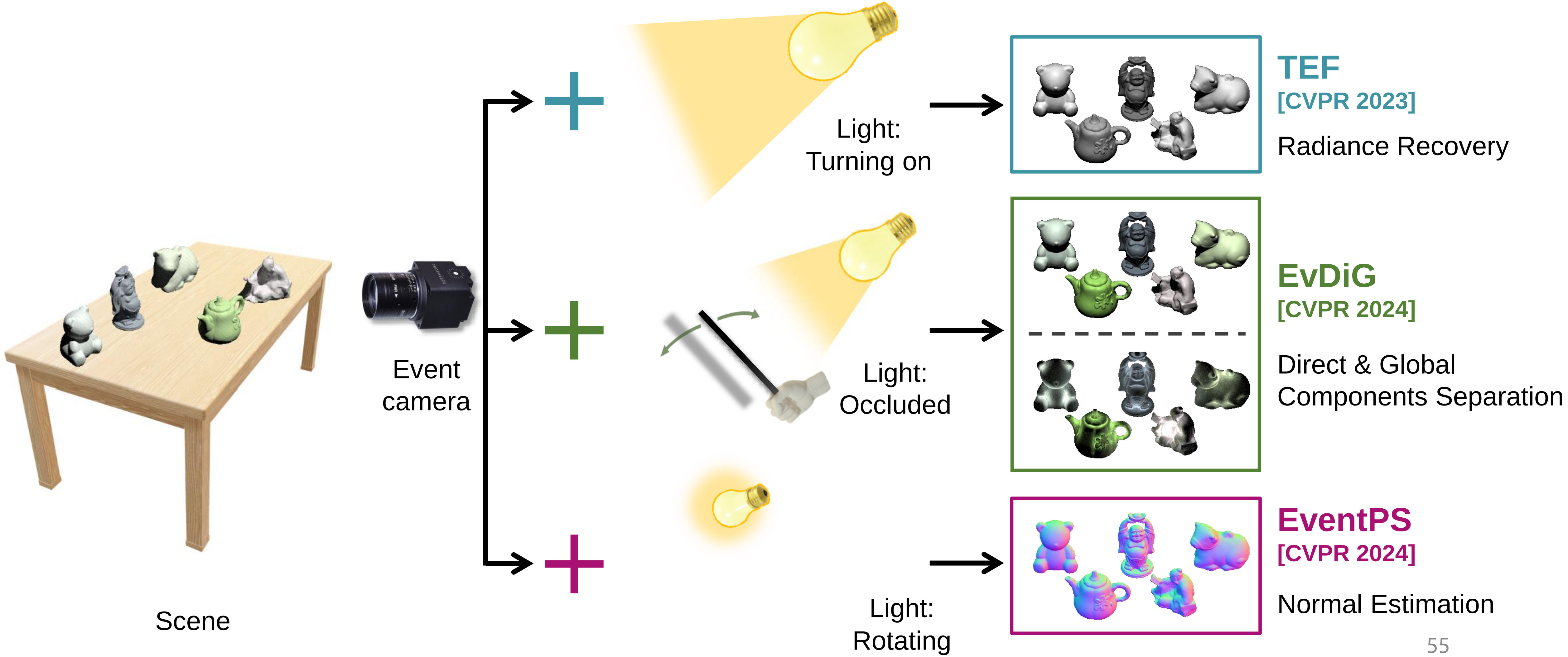
* SF-pattern-classic/SF-pattern-deep/reference use projector to generate the high frequency illumination patterns (cannot work in outdoor with sunlight)

Separation results on real-world dynamic scenes

Event Photometry



Event Photometry



EventPS: Real-Time Photometric Stereo Using an Event Camera

Bohan Yu[1] Jieji Ren[2] Jin Han[3] Feishi Wang[1] Jinxiu Liang[1] Boxin Shi[1*]

Oral / Best Paper, Runners-up

CVPR SEATTLE, WA JUNE 17-21, 2024

1 PEKING UNIVERSITY 1898
2 SHANGHAI JIAO TONG UNIVERSITY 1896
3 UTokyo

● Our Goal

Real-time PS using an event camera only

• Photometric Stereo (PS) Setup

Captured

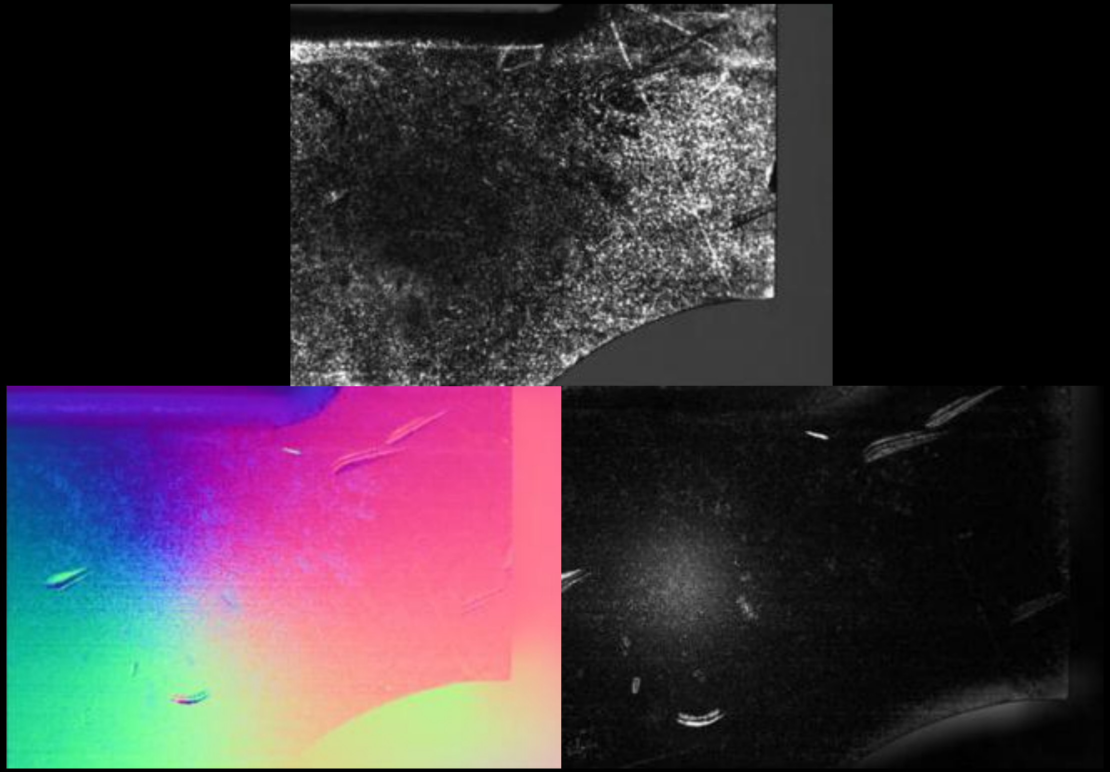
…

Normal map

[Shi et al. 10]

• Real-time PS Application

Face capturing [Andrew et al. 10]



Online product inspection [Michal et.al. 22]

Current PS Dilemma

Discrete dense – Moving light

Time consuming
Not real-time

Discrete sparse – Flashing light

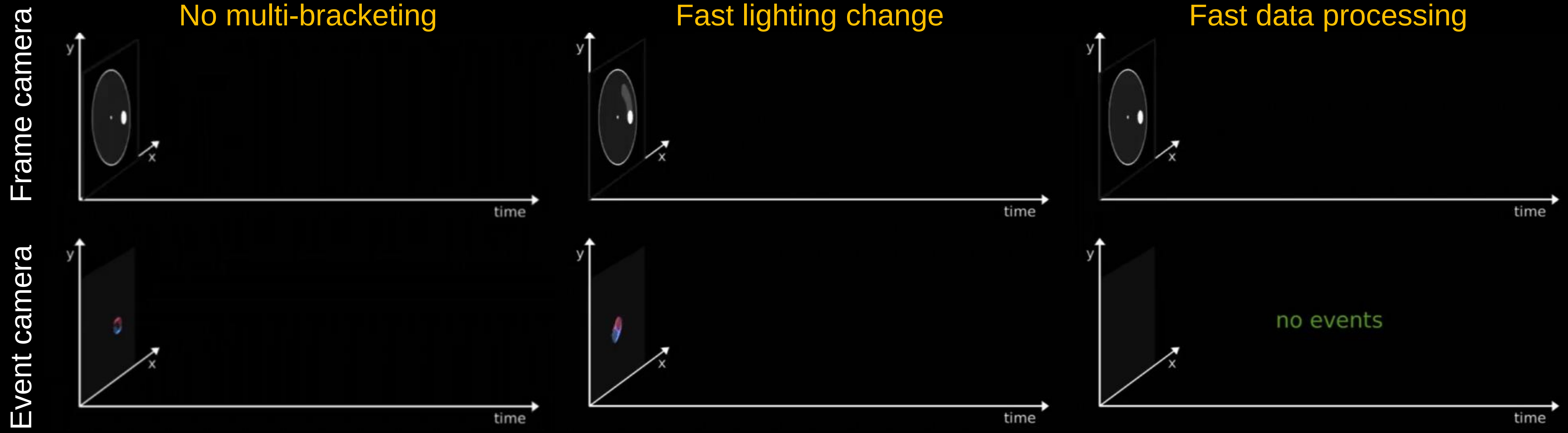
Few images, low dynamic range
Inaccurate

Event Camera

High dynamic range
No multi-bracketing

High temporal resolution
Fast lighting change

Sparse event representation
Fast data processing



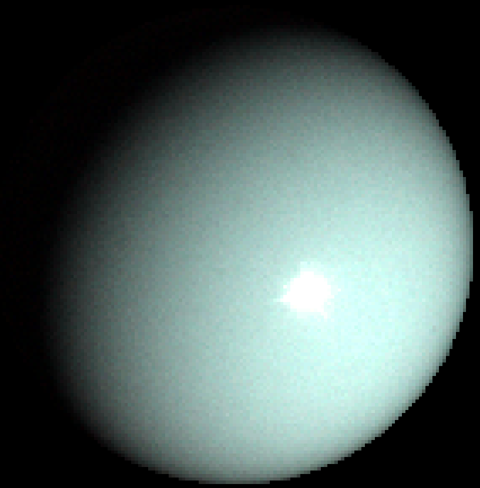
Video from: https://youtu.be/LauQ6LWTkxM

• Event Camera

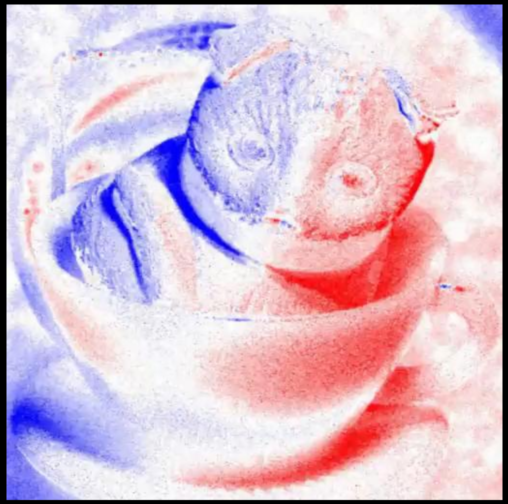
Motion related vision

*Video courtesy of Elias Mueggler

Photometric stereo



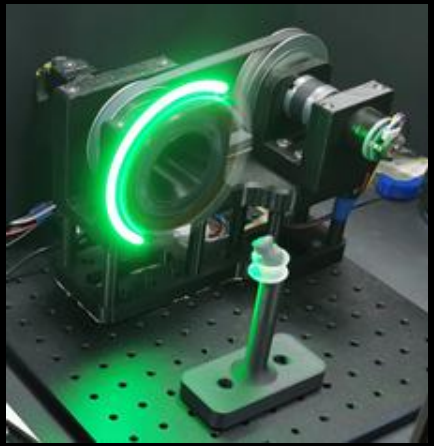
Frame camera



Event camera

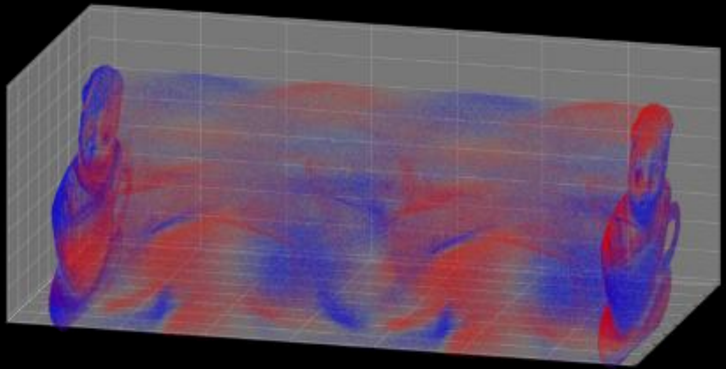
● EventPS Settings

How to illuminate the object for event camera?



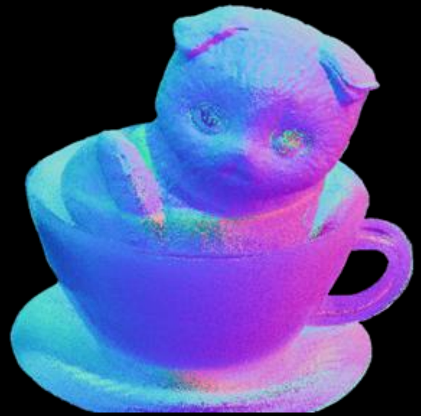
Continuously rotating light

~0.33s



Event camera
Fast, HDR, and sparse

?



Estimated surface normal

How to estimate surface normal without absolute intensity?

EventPS Model

Lambertian model

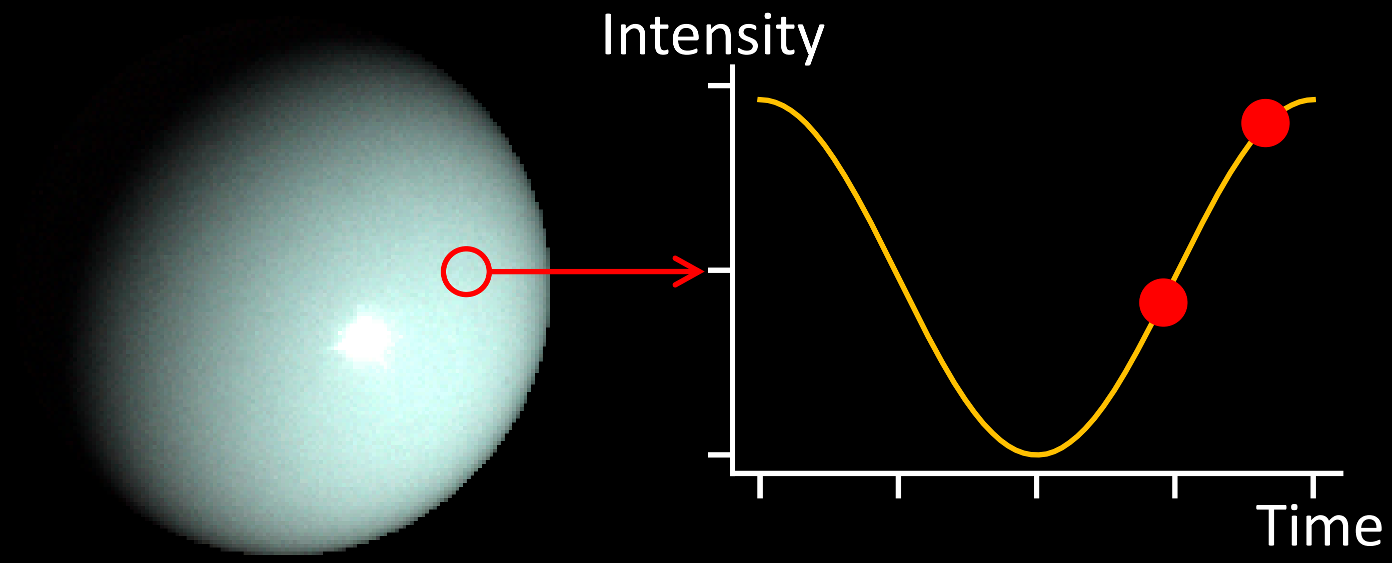
$$I_1 = \max\left(0, \alpha(\mathbf{n} \cdot \mathbf{l_1})\right)$$

Albedo
（Unknown）

Surface normal
（Target）

Lighting direction
（Calibrated）

EventPS Model



Lambertian model

$$I_1 = \max\big(0, \alpha(\mathbf{n} \cdot \mathbf{l_1})\big)$$

Albedo (Unknown)

Surface normal (Target)

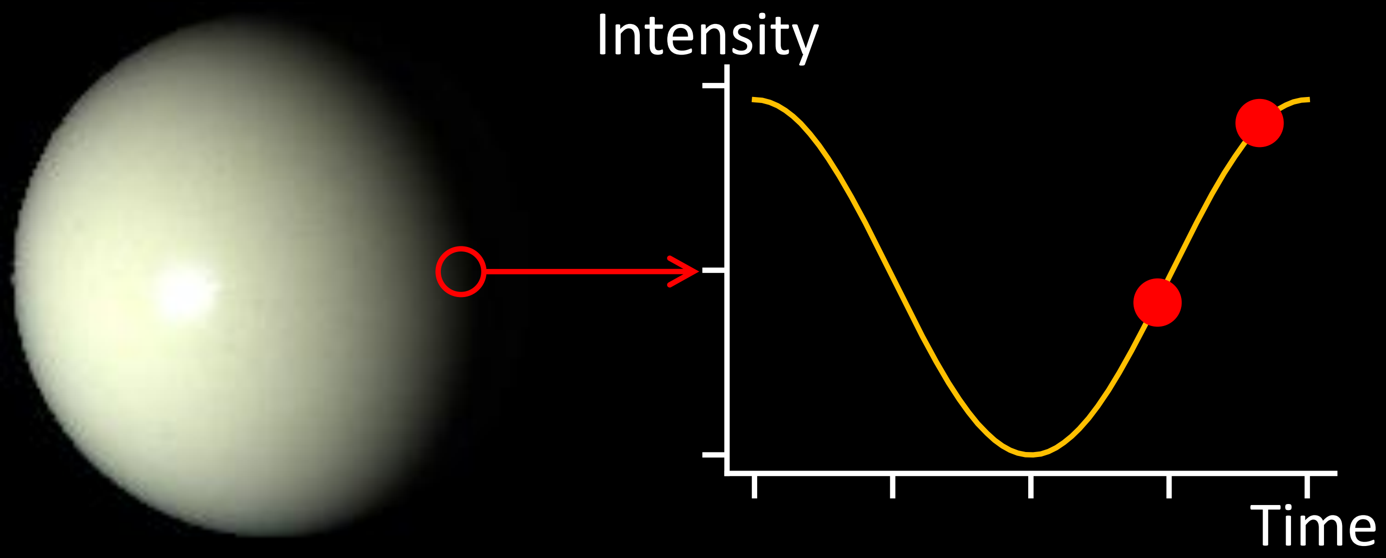
Lighting direction (Calibrated)

Event trigger model

$$I_1 = \sigma I_2$$

Event trigger threshold (Calibrated)

EventPS Model



$$\max\big(0, \alpha(\mathbf{n} \cdot \mathbf{l}_1)\big) = \sigma \max\big(0, \alpha(\mathbf{n} \cdot \mathbf{l}_2)\big)$$

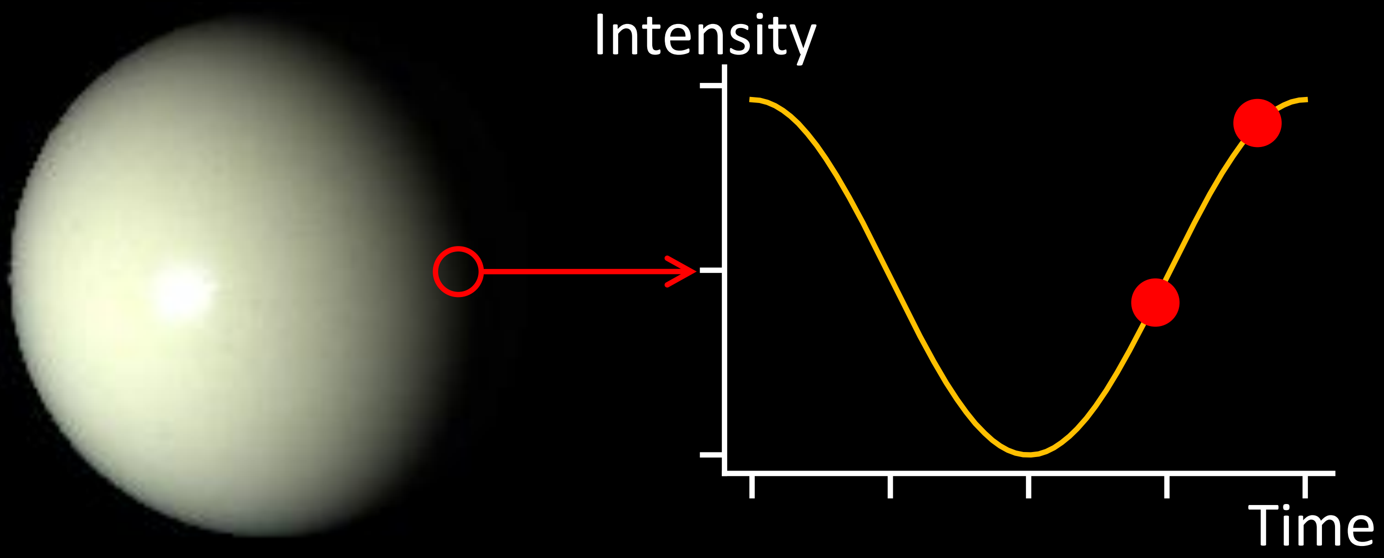
Albedo (Unknown)

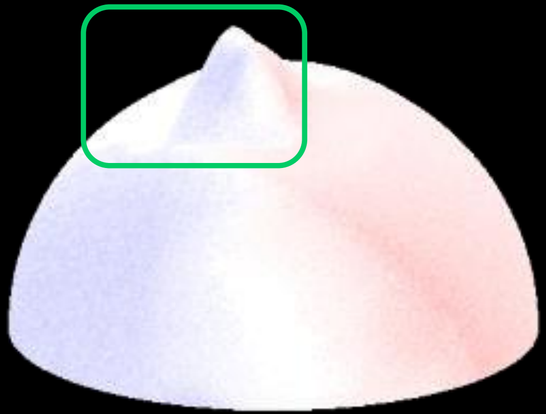
Surface normal (Target)

Lighting direction (Calibrated)

Event trigger threshold (Calibrated)

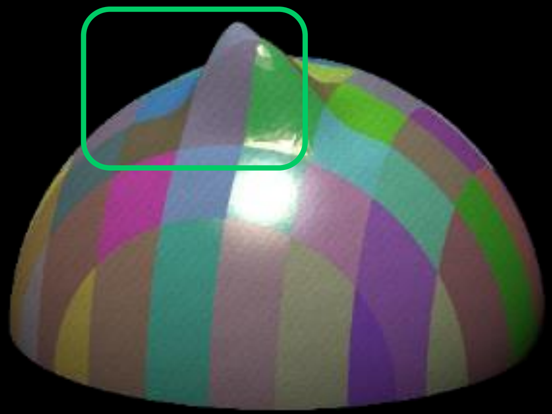
EventPS Model





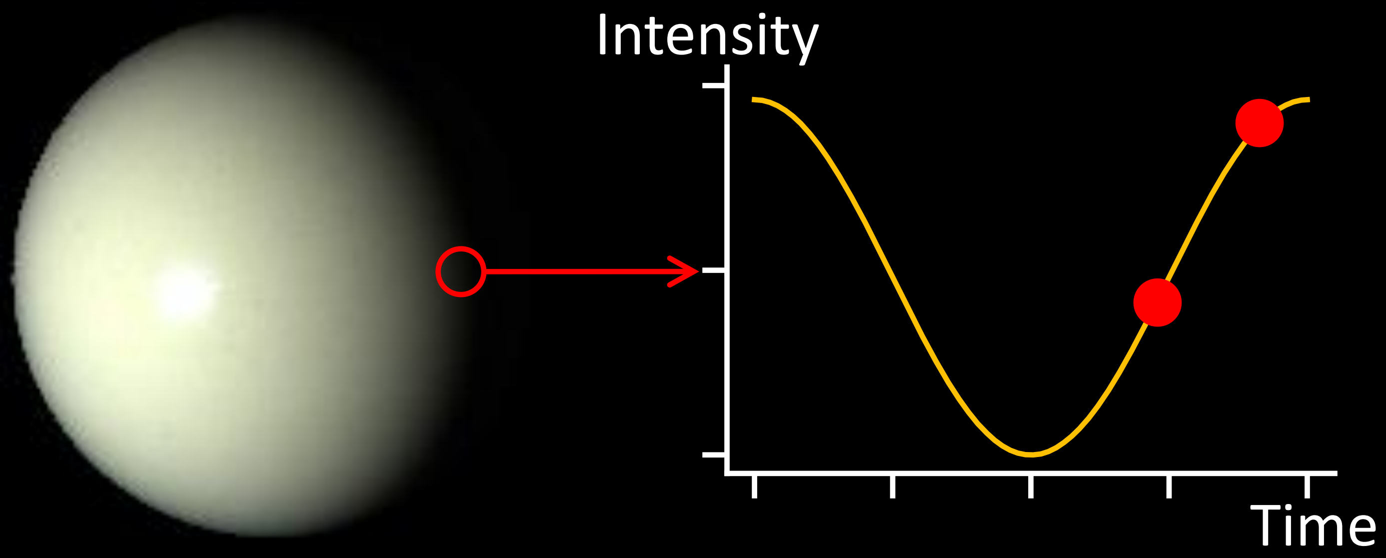
$$\max\left(0, \alpha(\mathbf{n} \bullet \mathbf{l}_1)\right) = \sigma \max\left(0, \alpha(\mathbf{n} \bullet \mathbf{l}_2)\right)$$

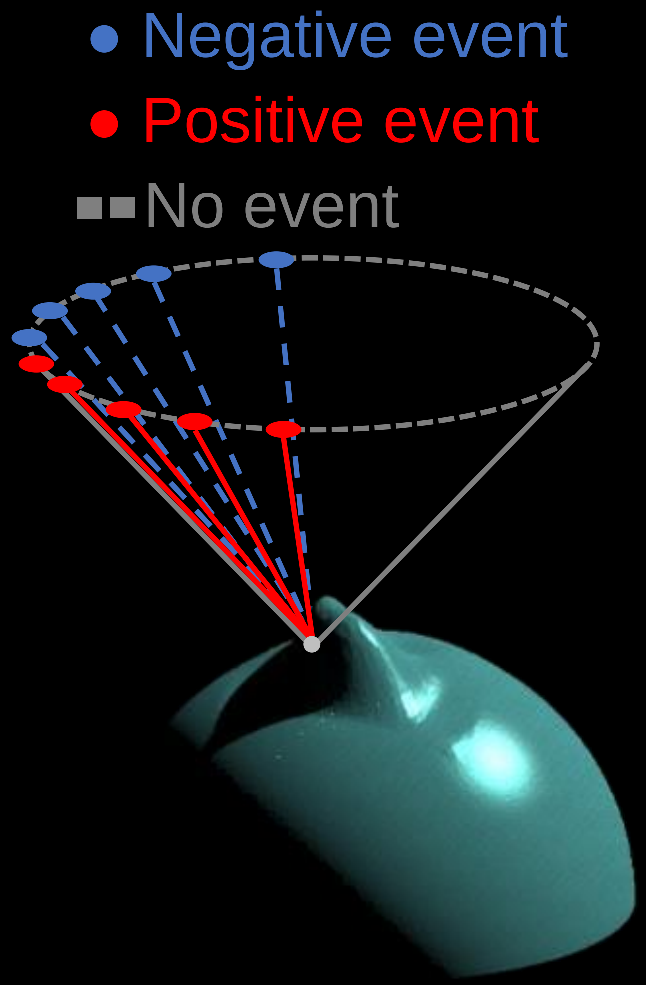
Observation 1: Albedo invariance



$$\max\left(0, (\mathbf{n} \bullet \mathbf{l}_1)\right) = \sigma \max\left(0, (\mathbf{n} \bullet \mathbf{l}_2)\right)$$

• EventPS Model



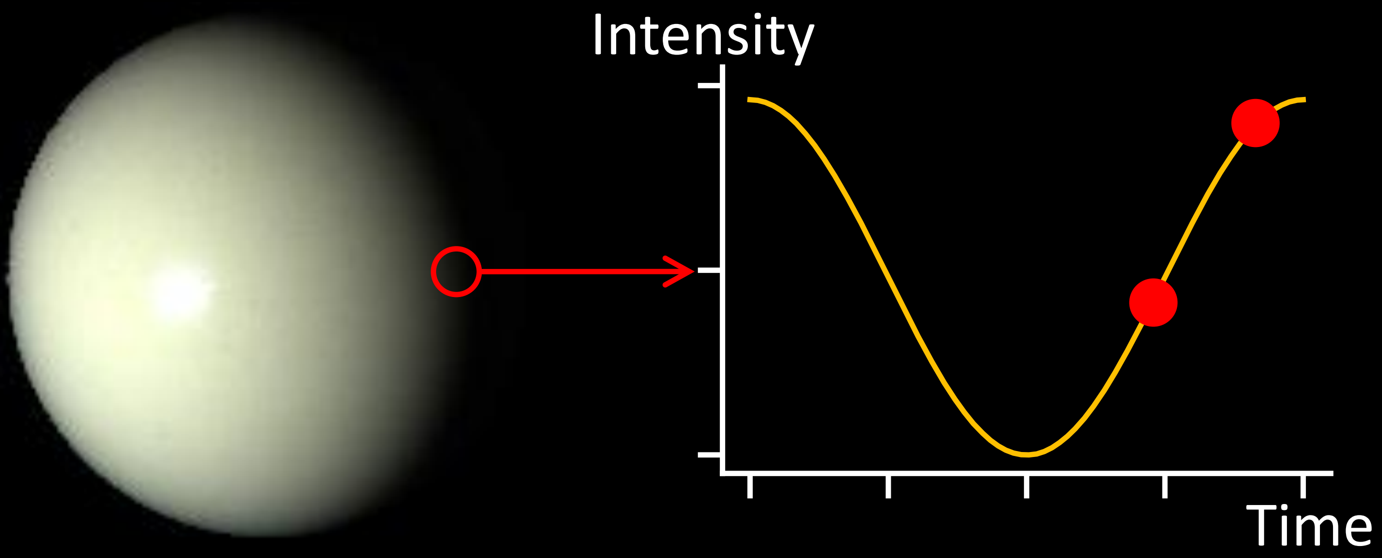


$$\max\left(0, (\mathbf{n} \cdot \mathbf{l}_1)\right) = \sigma \max\left(0, (\mathbf{n} \cdot \mathbf{l}_2)\right)$$

Observation 2: No events in attached shadow

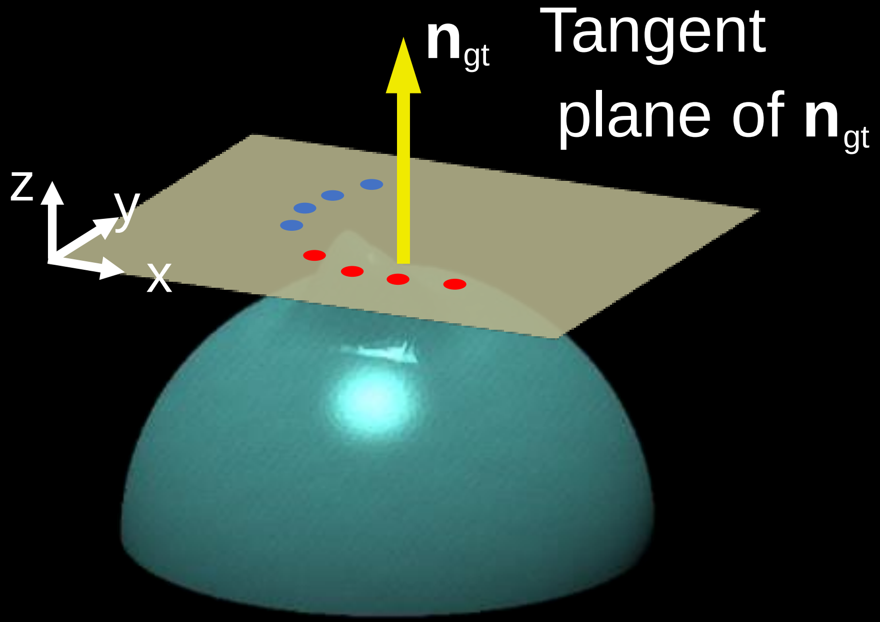
$$(\mathbf{n} \cdot \mathbf{l}_1) = \sigma (\mathbf{n} \cdot \mathbf{l}_2)$$

EventPS Model



Null space vectors from
Positive and Negative events

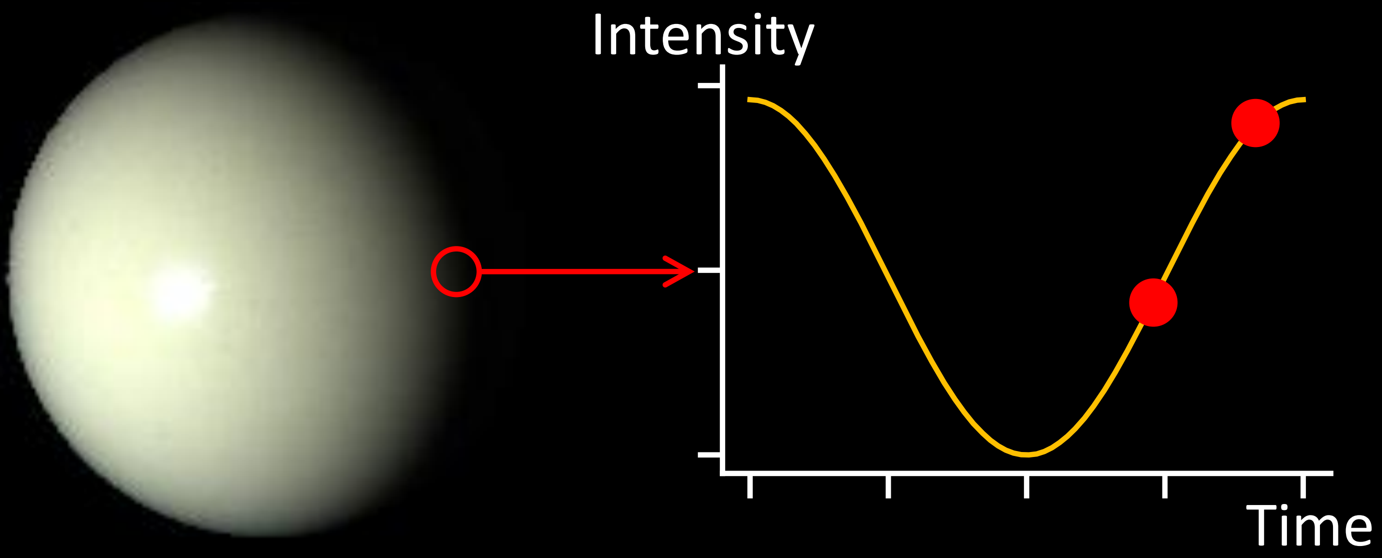
$$(\mathbf{n} \cdot \mathbf{l_1}) = \sigma(\mathbf{n} \cdot \mathbf{l_2})$$



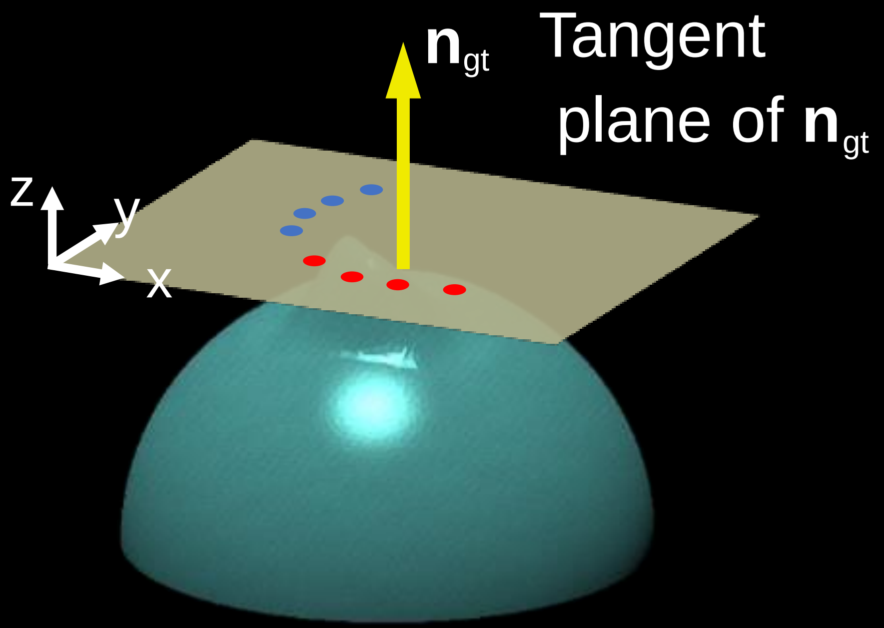
Observation 3: Null space vectors

$$(\mathbf{n} \cdot (\mathbf{l_1} - \sigma\mathbf{l_2})) = \mathbf{0}$$

EventPS Model



Null space vectors from
Positive and Negative events



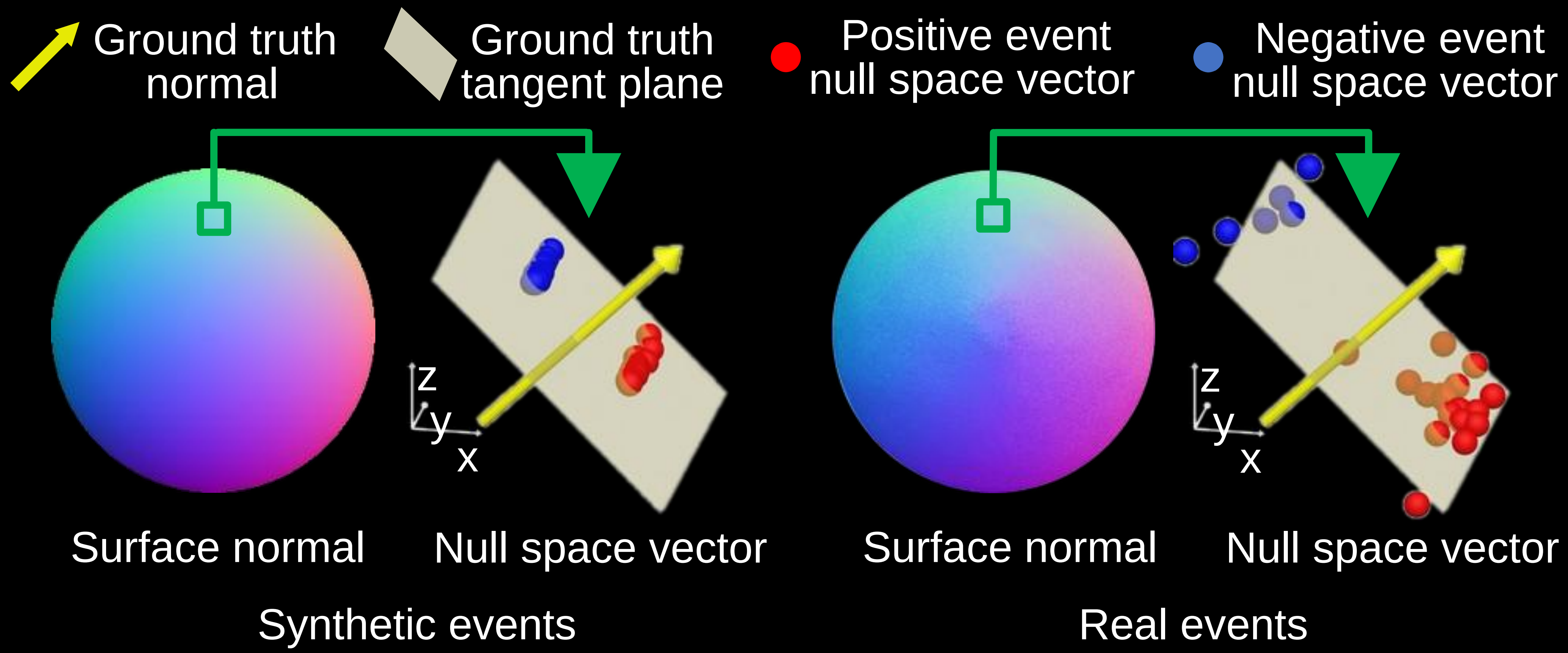
$$(\mathbf{n} \cdot \boxed{(\mathbf{l}_1 - \sigma \mathbf{l}_2)}) = \mathbf{0}$$

SVD solution

Null space vector

EventPS Model

Null space vectors perpendicular to surface normal

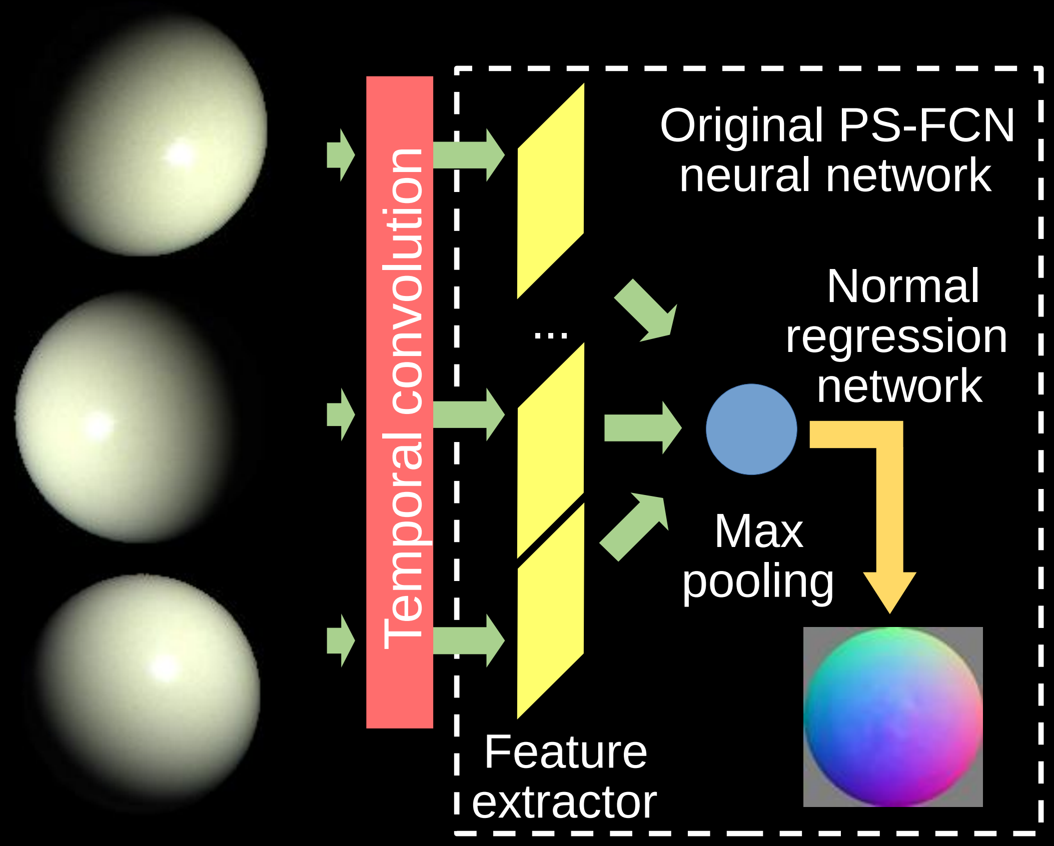


Validation Platform Components

• Deep Learning Variant

Deep learning variant to deal with generic BRDF:

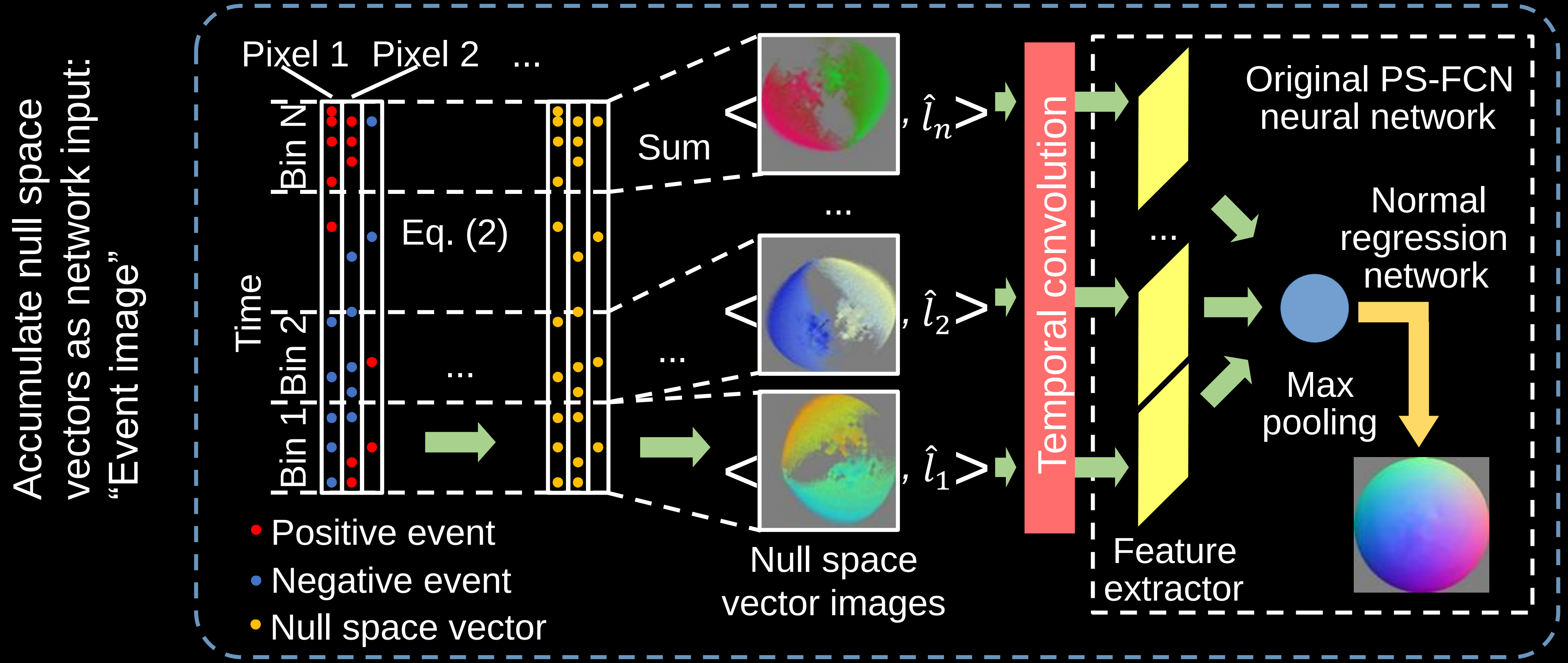
PS-FCN [Chen et al. 18]



Deep Learning Variant

Deep learning variant to deal with generic BRDF:

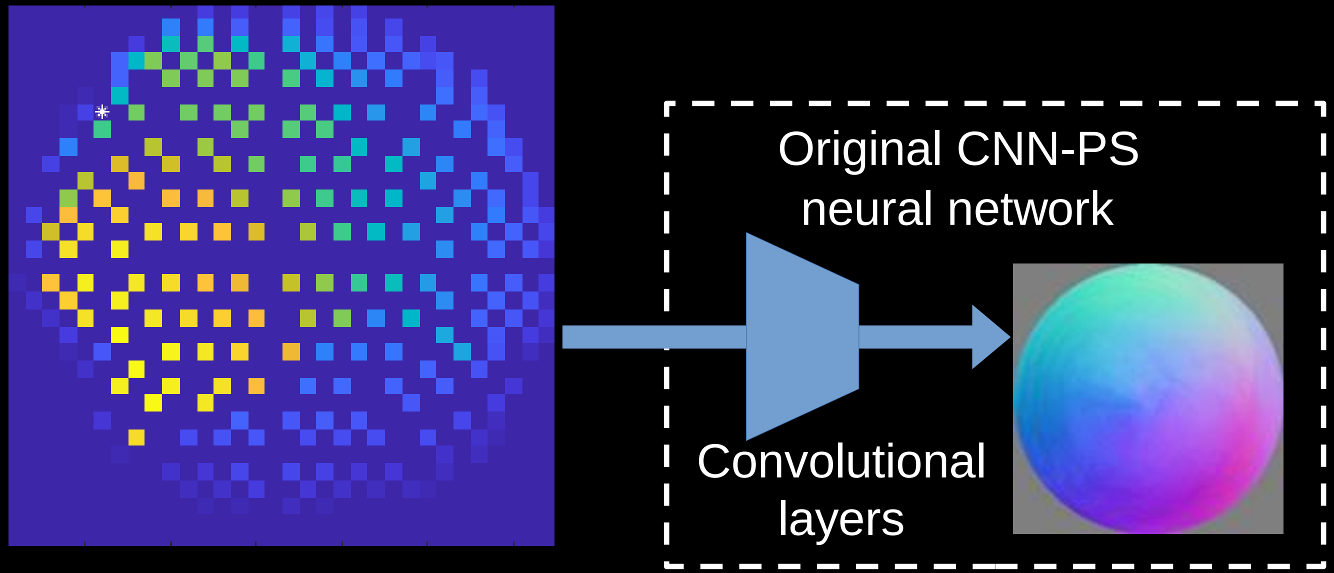
EventPS-FCN (Our extension)



Deep Learning Variant

Deep learning variant to deal with generic BRDF:

CNN-PS [Ikehata 18]



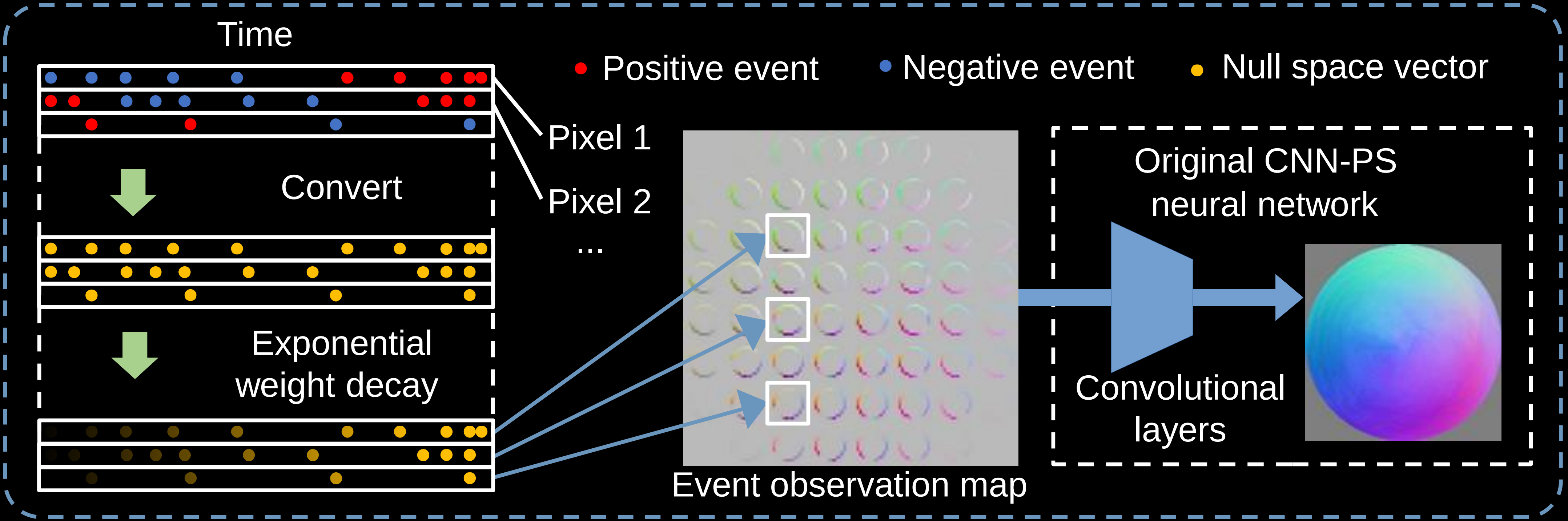
Observation map

• Deep Learning Variant

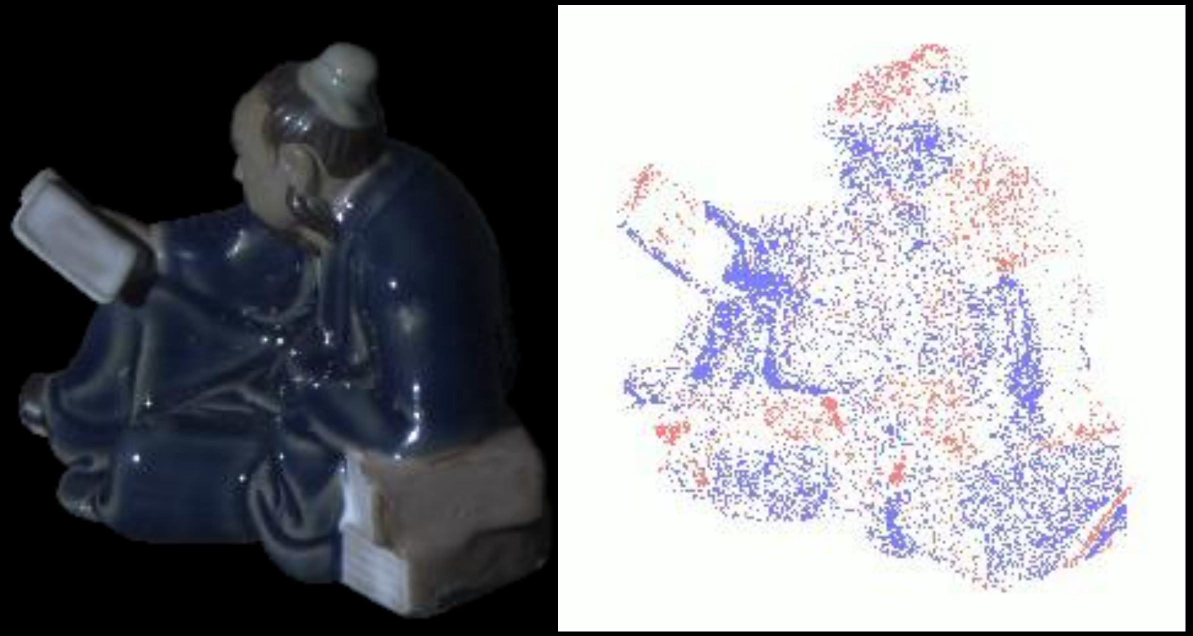
Deep learning variant to deal with generic BRDF:

EventPS-CNN (Our extension)

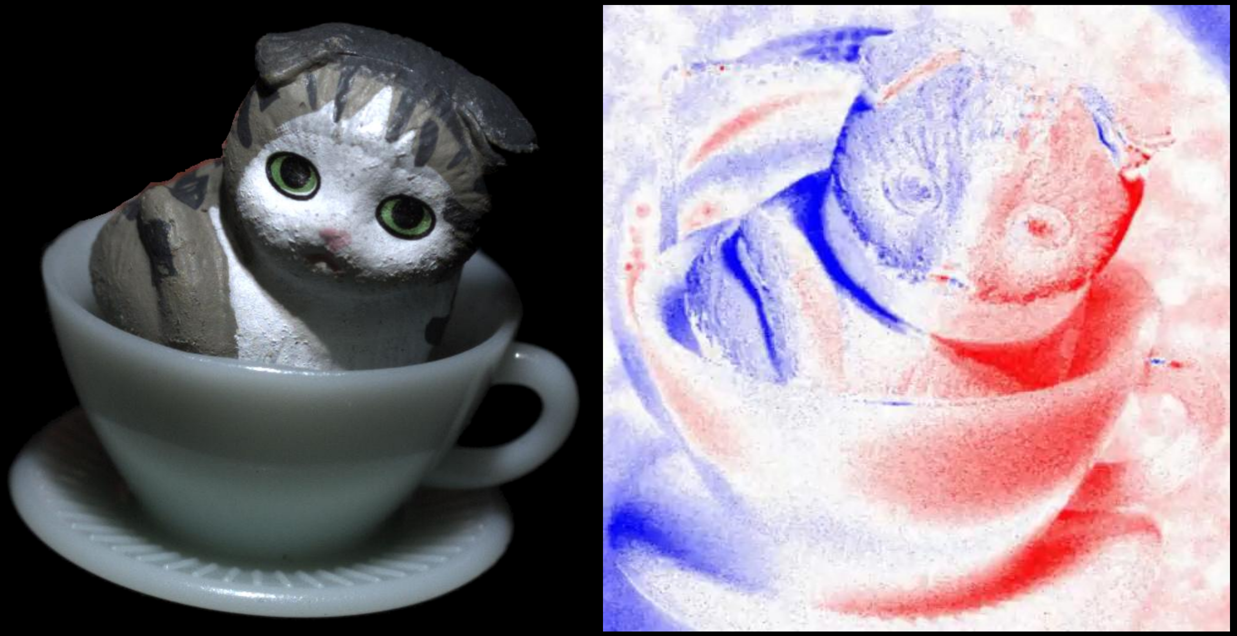
Build network input: “Event observation map”



● Experiment Settings



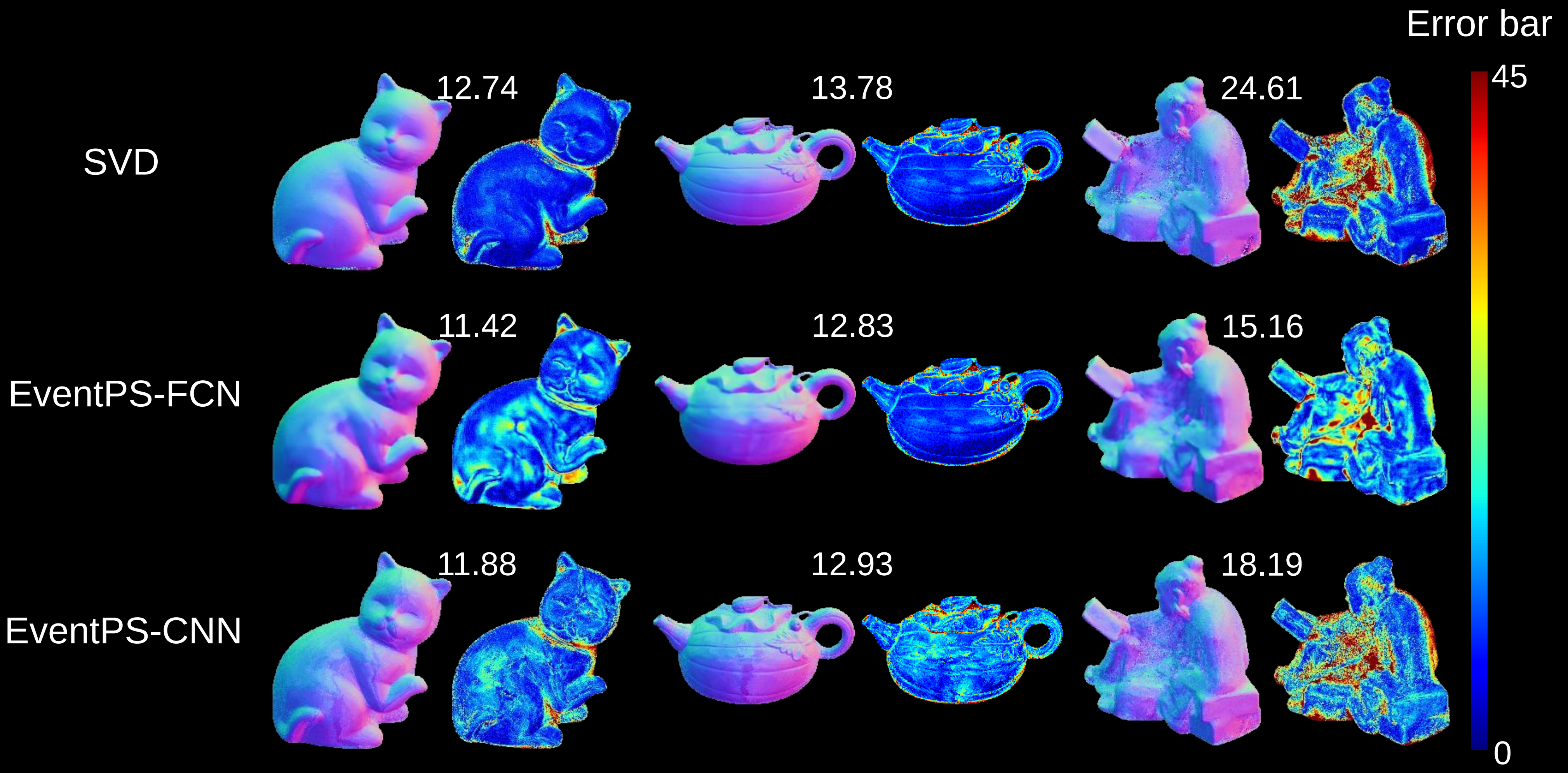
DiLiGenT [Shi et al. 19] semi-real
(Simulated from photos)



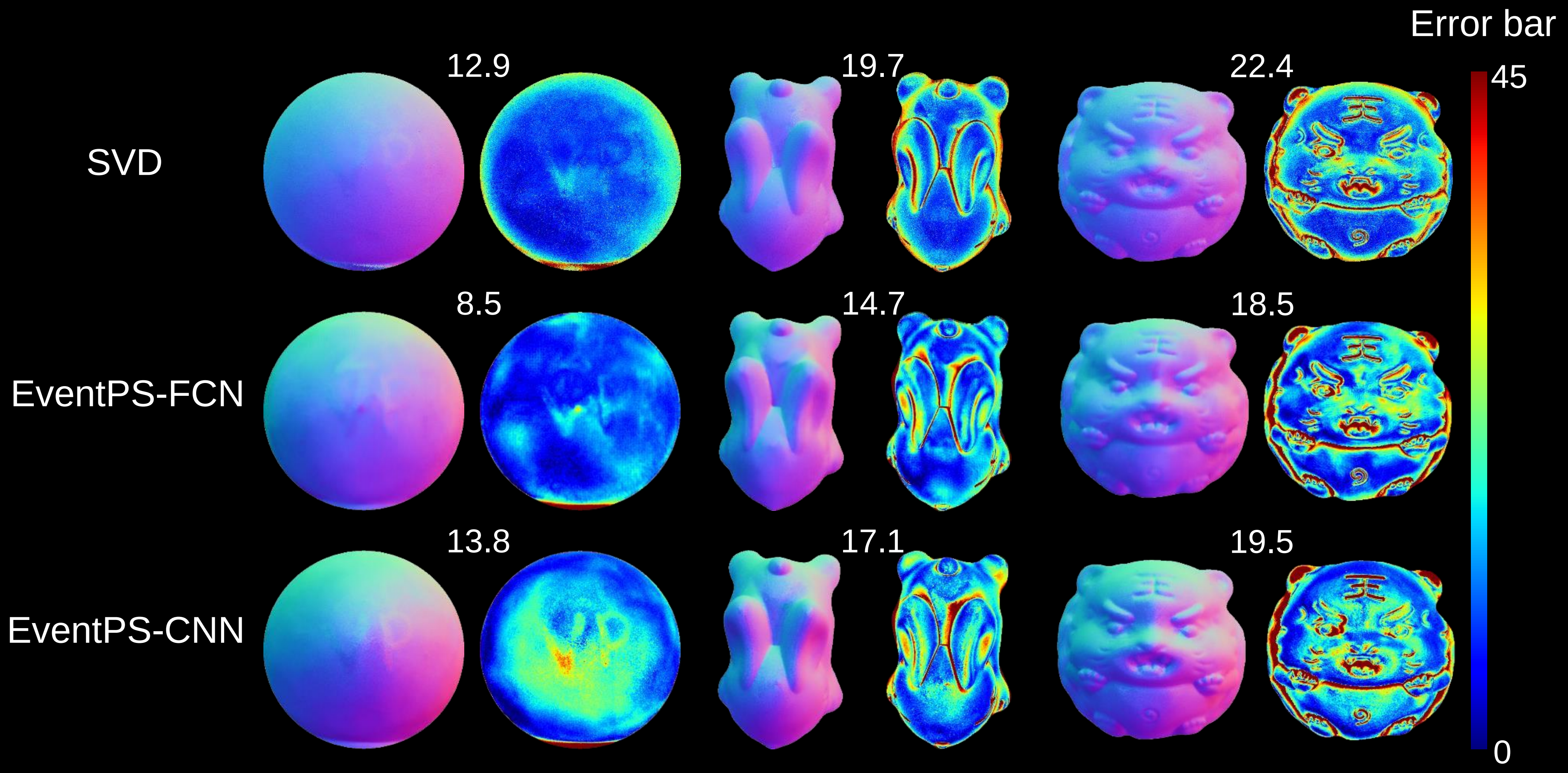
Real dataset
(Captured from Prophesee EVK4)

- **Tested Objects**

DiLiGenT Semi-real Dataset

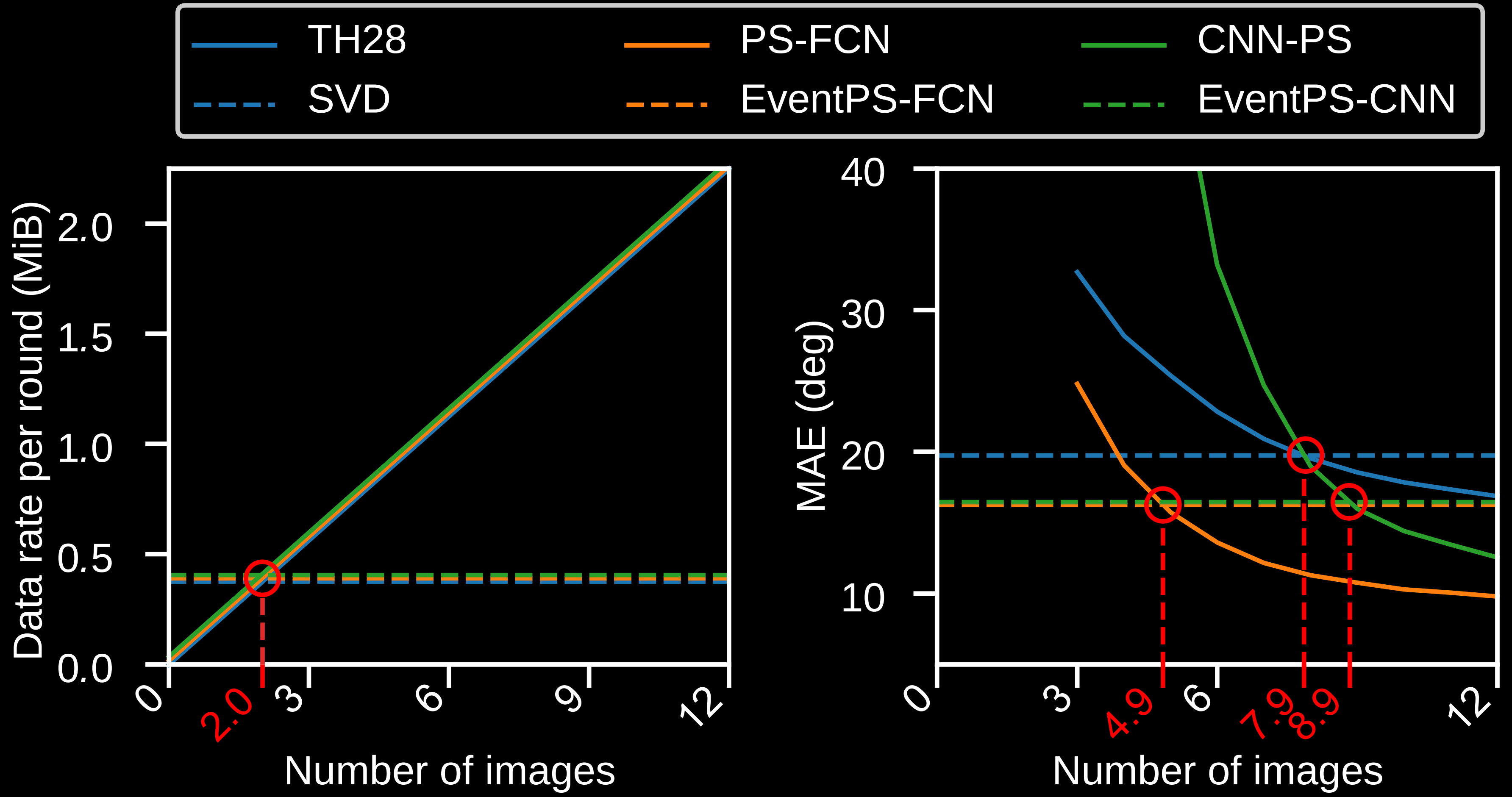


Real Dataset



Advantage

EventPS achieves higher accuracy with lower bandwidth



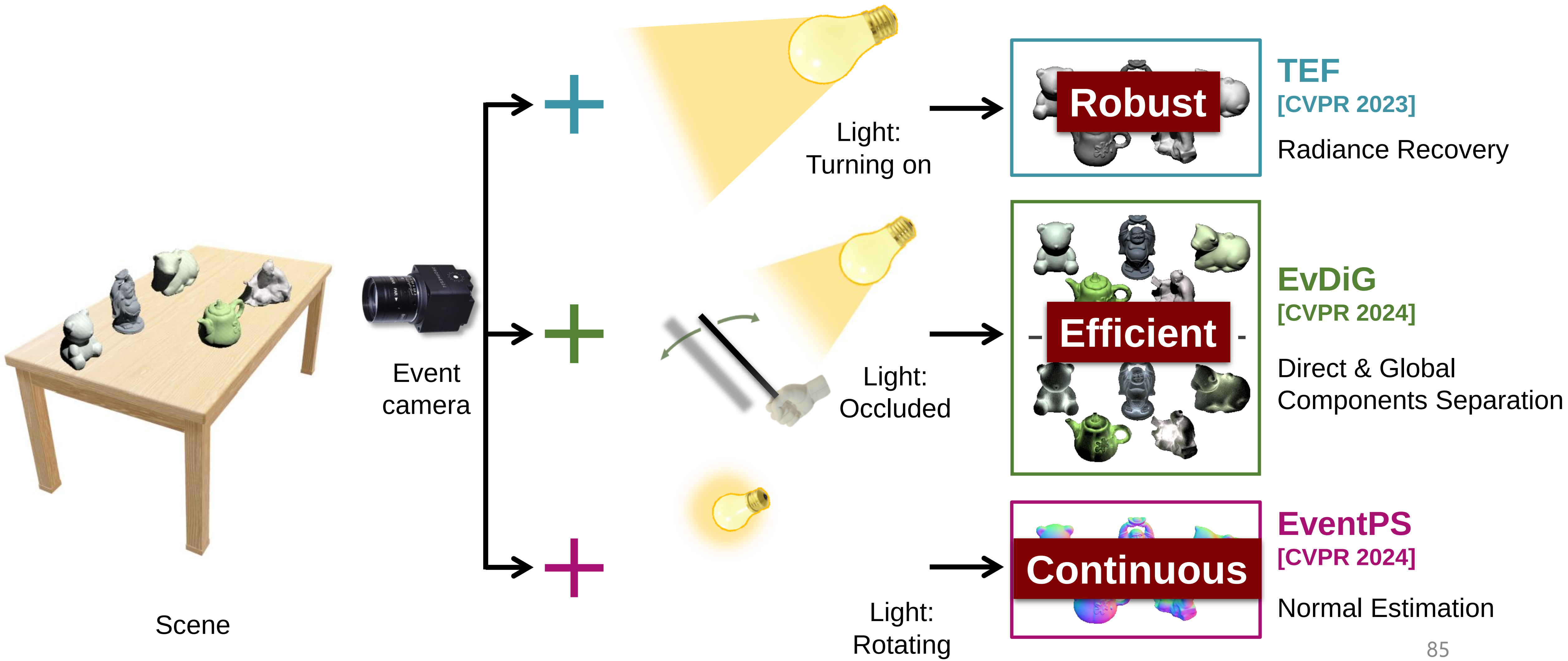
Low data consumption

High accuracy

● Be Real-time and Accurate

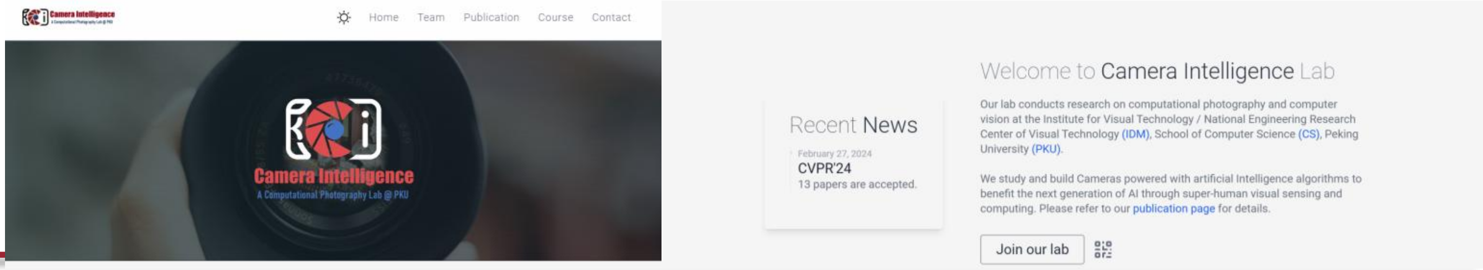
- Encoding photometric information to event timestamp
- Normal estimation tailored to event camera model
- SVD (GPU, OpenCL kernel) + deep learning (GPU)

Event Photometry

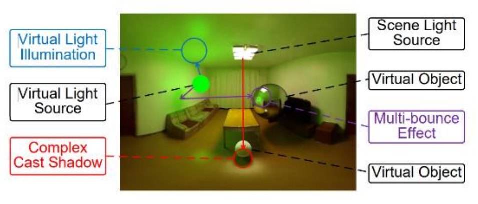


http://camera.pku.edu.cn

To know more …



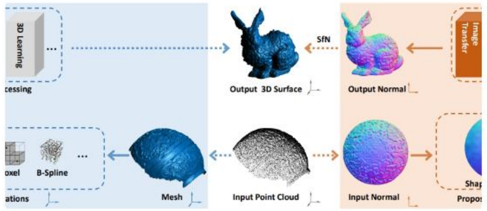
Select Year | TPAMI | Photometric ... | Clear | List View



[TPAMI23a] MILO: Multi-bounce inverse rendering for indoor scene with light-emitting objects

Bohan Yu, Siqi Yang, Xuanning Cui, Siyan Dong, Baoquan Chen, and Boxin Shi

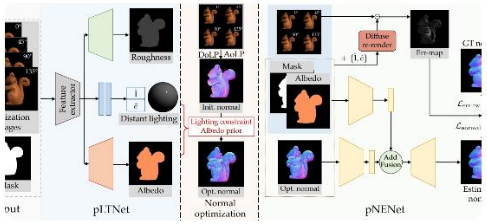
PDF | PHOTOMETRIC & 3D



[TPAMI23e] Surface geometry processing: An efficient normal-based detail representation

Wuyuan Xie, Miaohui Wang, Di Lin, Boxin Shi, and Jianmin Jiang

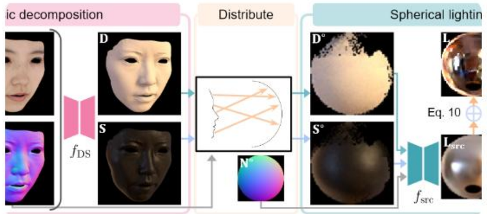
PDF | WEBSITE | PHOTOMETRIC & 3D



[TPAMI23f] Shape from polarization with distant lighting estimation

Youwei Lyu, Lingran Zhao, Si Li, and Boxin Shi

PDF | PHOTOMETRIC & 3D

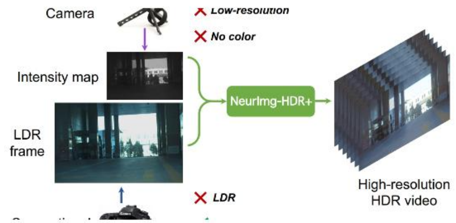


[TPAMI23h] SPLiT: Single portrait lighting estimation via a tetrad of face intrinsics

Fan Fei, Yean Cheng, Yongjie Zhu, Qian Zheng, Si Li, Gang Pan, and Boxin Shi

PDF | WEBSITE | PHOTOMETRIC & 3D

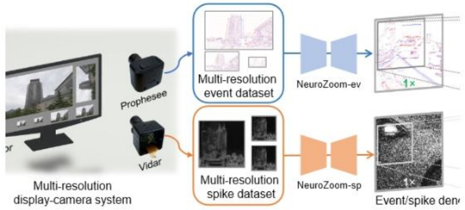
Select Year | TPAMI | Neuromorphi... | Clear | List View



[TPAMI23b] Hybrid high dynamic range imaging fusing neuromorphic and conventional images

Jin Han, Yixin Yang, Peiqi Duan, Chu Zhou, Lei Ma, Chao Xu, Tiejun Huang, Imari Sato, and Boxin Shi

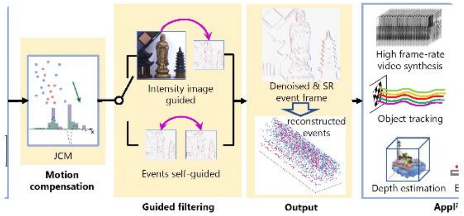
PDF | WEBSITE | NEUROMORPHIC CAMERA



[TPAMI23g] NeuroZoom: Denoising and super resolving neuromorphic events and spikes

Peiqi Duan, Yi Ma, Xinyu Zhou, Xinyu Shi, Zihao W. Wang, Tiejun Huang, and Boxin Shi

PDF | NEUROMORPHIC CAMERA



[TPAMI21b] Guided Event Filtering: Synergy between intensity images and neuromorphic events for high performance imaging

Peiqi Duan, Zihao W. Wang, Boxin Shi, Oliver Cossairt, Tiejun Huang, and Aggelos K. Katsaggelos

PDF | WEBSITE | NEUROMORPHIC CAMERA

Slides Courtesy

Jin Han
Final year PhD student
The University of Tokyo

hanjin1619@foxmail.com

Xinyu Zhou
Second year PhD student
Peking University

zhouxiny@pku.edu.cn

Bohan Yu
Third year PhD student
Peking University

ybh1998@pku.edu.cn

北京大学计算机学院
School of Computer Science

Camera Intelligence
A Computational Photography Lab @ PKU
http://camera.pku.edu.cn

Thank You
Q&A

Boxin Shi
shiboxin@pku.edu.cn

2024.06.27 @ SNL2024, Tokyo, Japan

* Courtesy of APSIPA Distinguished Lecturer Program