https://yunfanlu.github.io/

Event Cameras Meet Computational Imaging

For Fast Motion, High Resolution, Extreme Illumination, and Faithful Color

Presenter: **Yunfan Lu**

Supervisor: Prof. Hui Xiong

Institution: The Hong Kong University of Science and Technology (GZ)

2026.05.28

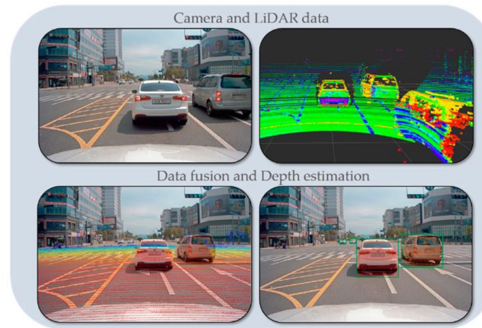
Part I. Motivation and Perspective

Why Rebuild Imaging with Events?

Why Rethink Visual Sensing?

Biological vision remains robust where frame-based sensing often fails.
Especially under fast motion, low light, and extreme illumination changes.

The gap may begin at sensing.



Bird vs. drone
A bird can fly fast and avoid obstacles, but a modern drone with many sensors still cannot fly as safely.

Human driver vs. autonomous car
A human uses only two eyes, yet a car with many sensors still struggles in complex scenes.

Limits of Frame-Based Sensing

1. Motion is compressed into frames

Fast dynamics are integrated over exposure time, causing blur, distortion, and temporal loss.

2. Light is compressed into limited exposure and bit depth

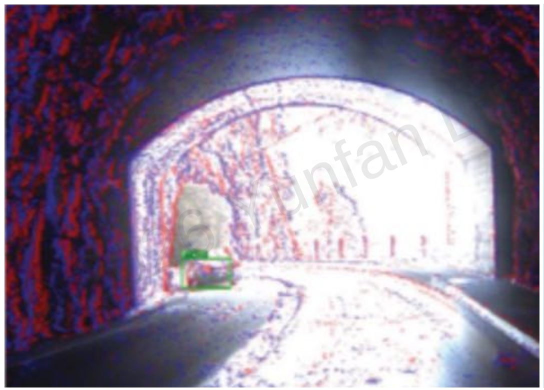
Mixed lighting leads to overexposure, underexposure, and loss of detail.

3. Image quality is constrained by sensor trade-offs

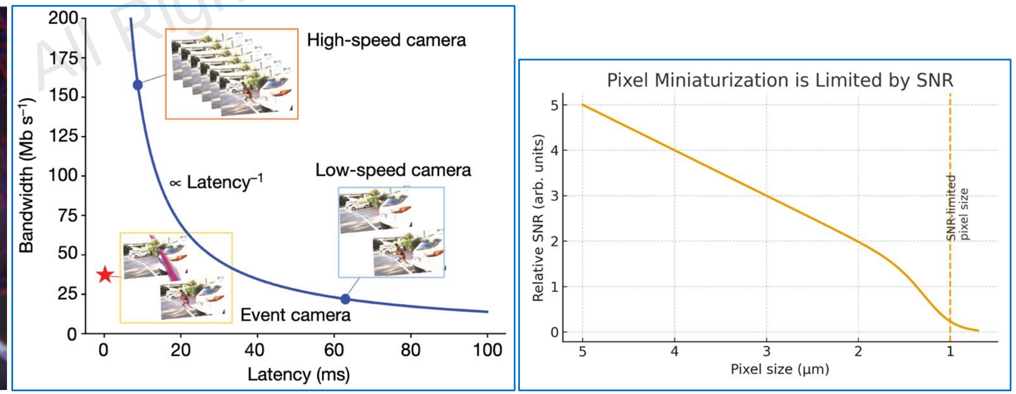
Higher frame rate, smaller pixels, and lower light often reduce SNR and reliability.

What is the root cause of these failures?

(a) **Fast motion**
Frame exposure collapses rapid motion into blur.



(b) **Extreme illumination**
A single exposure cannot capture both dark and bright regions.



(c) **Full-frame sensing creates sensor trade-offs**
Higher speed and smaller pixels often come at the cost of bandwidth and signal quality.

https://yunfanlu.github.io/

Frame Paradigm Has Barely Changed

Root Cause: Synchronous Exposure + Full-Frame Readout

This frame paradigm collapses continuous motion and light changes into discrete frames.

Can we sense the world without forcing it into frames?

Metal Photochemical Photography,
Obscura Box Camera,
Joseph Nicephore, 1826

The Birth of Camera

86 years

Photochemical Photography,
UrLeica Camera,
Oskar Barnack, 1912

Portable Camera

63 years

Digital Photography
Kodak Camera,
Steve Sasson, 1975

Digitization of Camera

Upgrade ✅
1. Semiconductor process
2. Algorithm

Maintained ❌
Exposure mechanism

50 years

Consumer Camera,
and Hasselblad, 2025

Consumer-Grade Camera

What is still needed
1. Higher temporal resolution
2. Wider dynamic range
3. Lower-power sensing

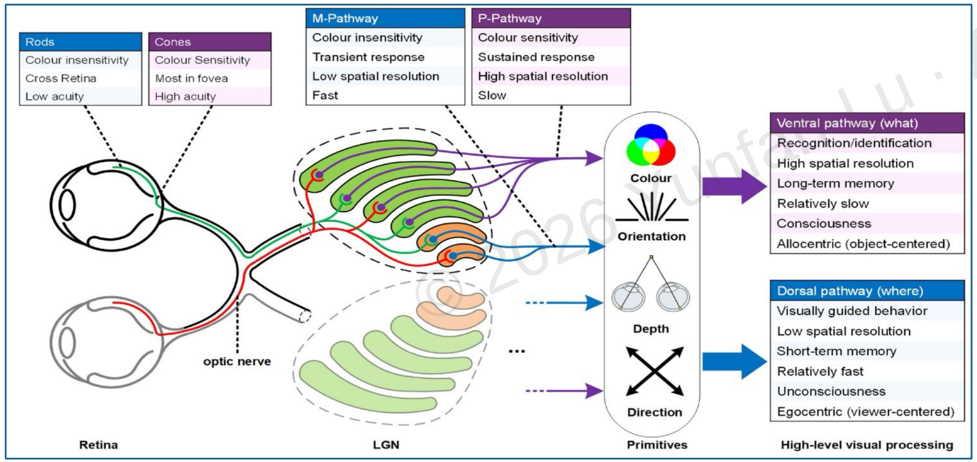
Vision comparable to biology 👀

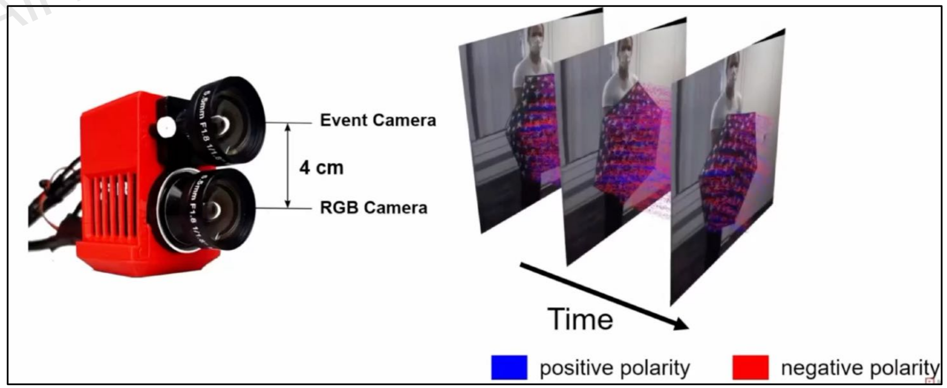
Sensing Without Frames: Event Cameras

Event cameras asynchronously (~~Synchronous Exposure~~) record brightness changes (~~Full-Frame Readout~~).

- Low latency (<1ms)
- High dynamic range (>120dB)
- Low power (~36 μW)
- Low bandwidth demand (94% reduction)

How far can event cameras go for imaging?





[1] Gehrig D, Scaramuzza D. Low-latency automotive vision with event cameras[J]. Nature, 2024, 629(8014): 1034-1040.
[2] Yang, Zheyu, et al. "A vision chip with complementary pathways for open-world sensing." Nature 629.8014 (2024): 1027-1033.

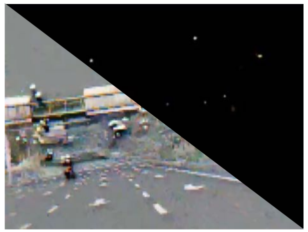
https://yunfanlu.github.io/

Rebuilding Imaging with Events

A three-year effort to position events as a key modality for reshaping the imaging pipeline, with gains in time, space, dynamic range (DR), and color.

Next, I will walk through this effort from these directions.

Re-Exposure
DR, IJCV 2025 *



Low Light Enhancement
DR, CVPR 2024

Rolling Shutter Correction
Time, IEEE TVCG 2024 *

Deblurring
Time ECCV 2024 *

Dehaze & Diffision Model
DR, CVPR 2026 *

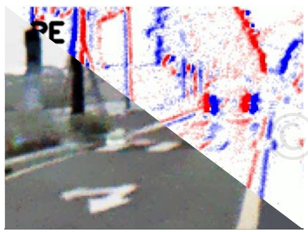
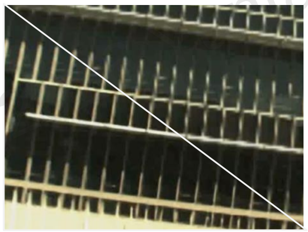


Image Reconstruction
Space, ECCV 2024



Super-Resolution
Space, CVPR 2023 *

Frame Interpolation
Time, Space, IJCV 2025 *

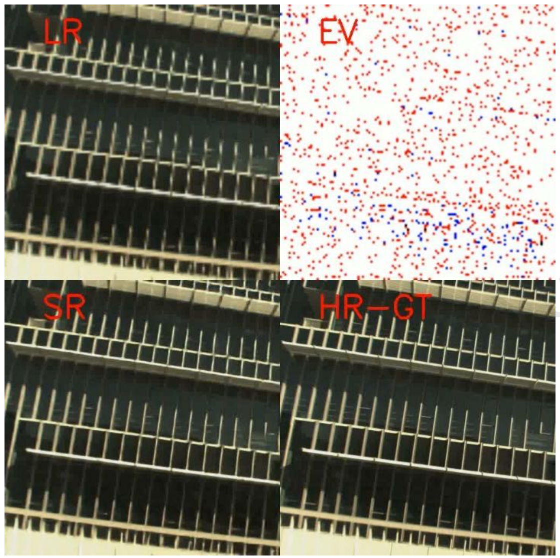
Color Correction
Color, ICLR 2025 *

Demosaic & Denoise
Color, CVPR 2025

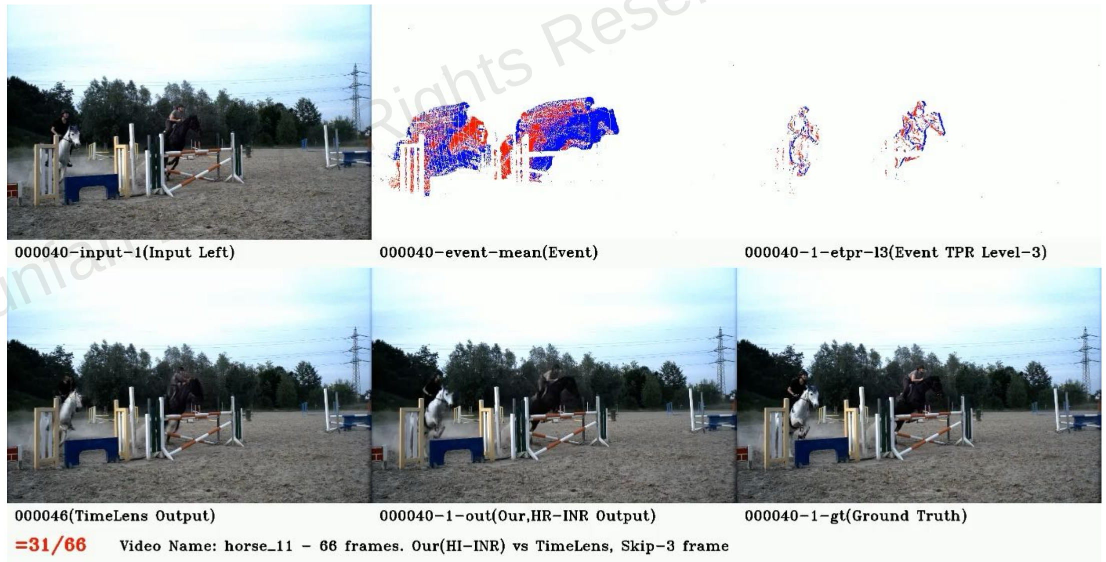
** denotes first author.*

Spatial-Temporal Reconstruction

- Step 1 – Events recover spatial detail lost in low-resolution frames. CVPR 2023, IEEE-TVCG 2025
- Step 2 – Toward a unified space-time representation. ECCV 2024
- Step 3 – Events enable continuous-time video reconstruction beyond fixed frame sampling. IJCV 2025
- Key Insight: Events provide the missing temporal structure needed to reconstruct content jointly in space and time.



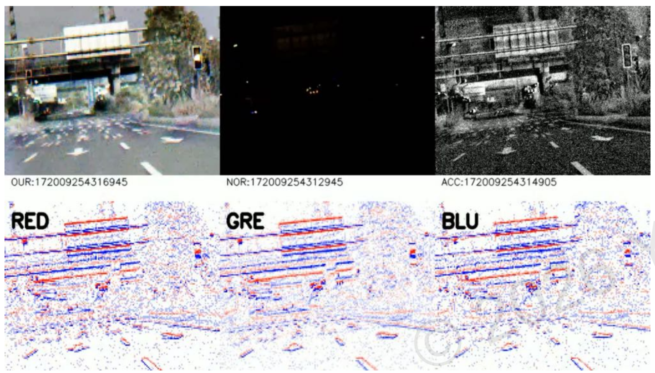
(a) Events guided Video Super-resolution, CVPR 2023



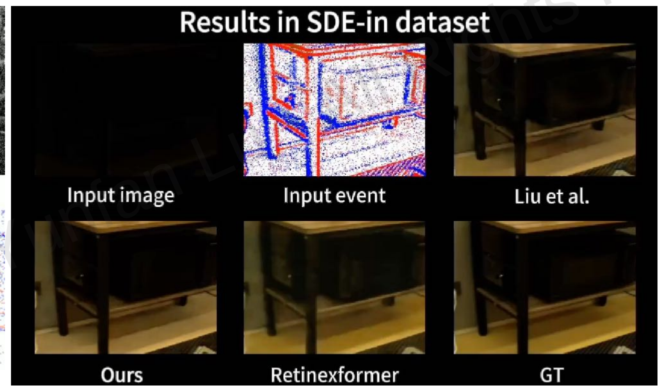
(b) High-Frame-Rate Video Reconstruction, ECCV 2024, IJCV 2025

Dynamic Range Recovery

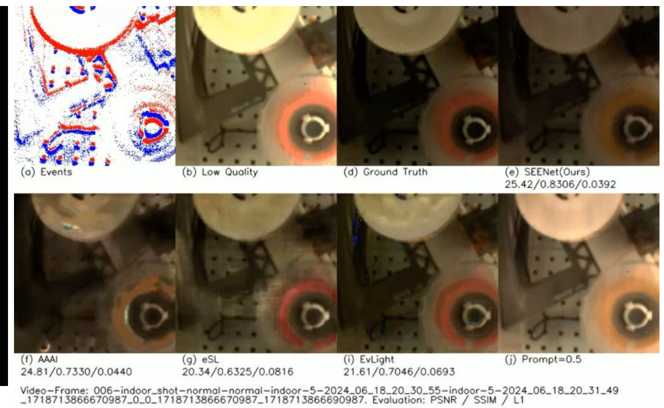
- Step 1 – Events recover scene information under low illumination. ECCV 2024
- Step 2 – Events provide missing cues to restore low-light frames when photon-limited images become unreliable. CVPR 2024, IEEE-TPAMI 2025
- Step 3 – Events enable exposure-recovery pipelines that adapt imaging beyond fixed exposure. IJCV 2025
- Key Insight: Dynamic range recovery is fundamentally a sensing problem, and events provide a widely signal to recover what frame exposure discards.



(a) Video reconstruction directly from events streaming, ECCV 24



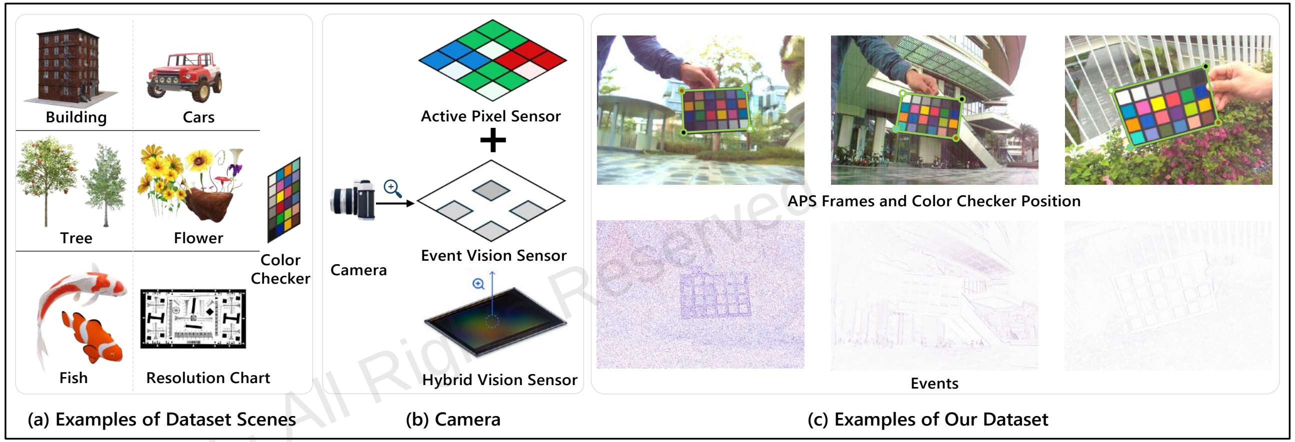
(b) Low light video enhancement from events and frames, CVPR 24 TPAMI 25



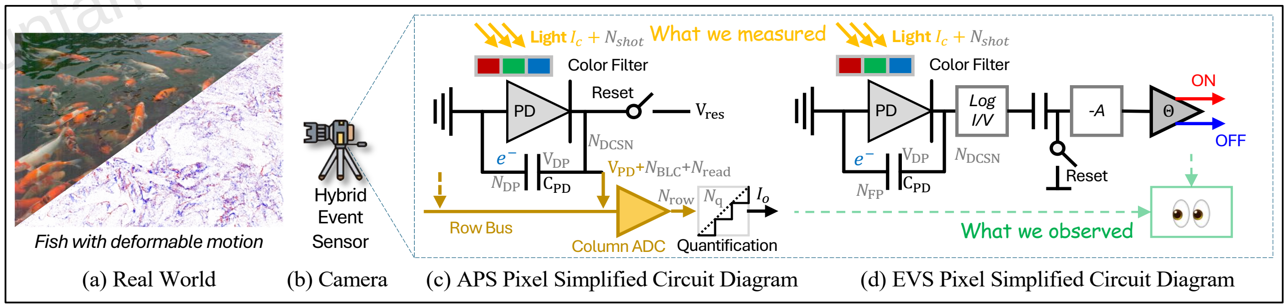
(c) Low light video enhancement from events and frames, IJCV 25

RAW & Color

- Step 1 – RAW-level event-frame fusion opens a new route to event-aware image signal processing (ISP). ICLR 2025
- Step 2 – Hybrid sensor modeling connects sensing physics, calibration, and learning-based reconstruction. Under Review
- Key Insight: Events are not only temporal signals. When combined with RAW measurements and sensor models, they provide physical cues for motion, illumination, and reflectance, opening a new path toward unified color imaging.



(a) RGB-Event ISP Dataset and Benchmark, ICLR 2025



(b) Hybrid Event Vision Sensor Simulator, ECCV 2026

Publication - *1/2*

Space–Time

- Yunfan LU, Zipeng Wang, Yusheng Wang, Hui Xiong; Continuous Space-Time Video Super-Resolution via Event Camera, IJCV 2025 CCF-A, TH-CPL-A
- Yunfan LU, Guoqiang Liang, YiranShen, Lin Wang; Self-supervised learning of event-guided video frame interpolation for rolling shutter frames, IEEE TVCG 2025, CCF-A, TH-CPL-A
- Yunfan LU, Guoqiang Liang, Yusheng Wang, Lin Wang, Hui Xiong; UniINR: Event-guided unified rolling shutter correction, deblurring, and interpolation, ECCV 2024, CCF-B, TH-CPL-A
- Yunfan LU, Zipeng Wang, Minjie Liu, Hongjian Wang, Lin Wang; Learning Spatial-Temporal Implicit Neural Representations for Event-Guided Video Super-Resolution, CVPR 2023, CCF-A, TH-CPL-A

Dynamic Range

- Yunfan LU, Xiaogang Xu, Hao LU, Yanlin Qian, Bin Yang, Junyi Li, Qianyi Cai, Weiyu Guo, Hui Xiong; SEE: See Everything Every Time - Adaptive Brightness Adjustment for Broad Light Range Images via Events, IJCV 2025, CCF-A, TH-CPL-A
- Kanghao Chen*, Guoqiang Liang*, Yunfan LU *, Hangyu Li, Lin Wang; Evlight++: Low-light Video Enhancement with an Event Camera: A Large-scale Real-world Dataset, Novel Method, and More, IEEE T-PAMI 2025 * These authors contributed equally. CCF-A, TH-CPL-A
- Ling Wang*, Yunfan LU *, Wenzong Ma, Huizai Yao, Pengteng Li, Hui Xiong; From Events to Clarity: The Event-Guided Diffusion Framework for Dehazing, CVPR2026 * These authors contributed equally.

Publication - 2/2

RAW & Color

- Yunfan LU, Yanlin Qian, Ziyang Rao, Junren Xiao, Liming Chen, Hui Xiong; RGB-Event ISP: The Dataset, Benchmark and Direction, ICLR 2025. TH-CPL-A

Sensor Modeling

- Yunfan LU, Nico Messikommer, Xiaogang Xu, Liming Chen, Yuhan Chen, Nikola Zubic, Davide Scaramuzza, Hui Xiong; Hybrid Event Frame Sensors: Modeling, Calibration, and Simulation, ECCV 2026

Survey

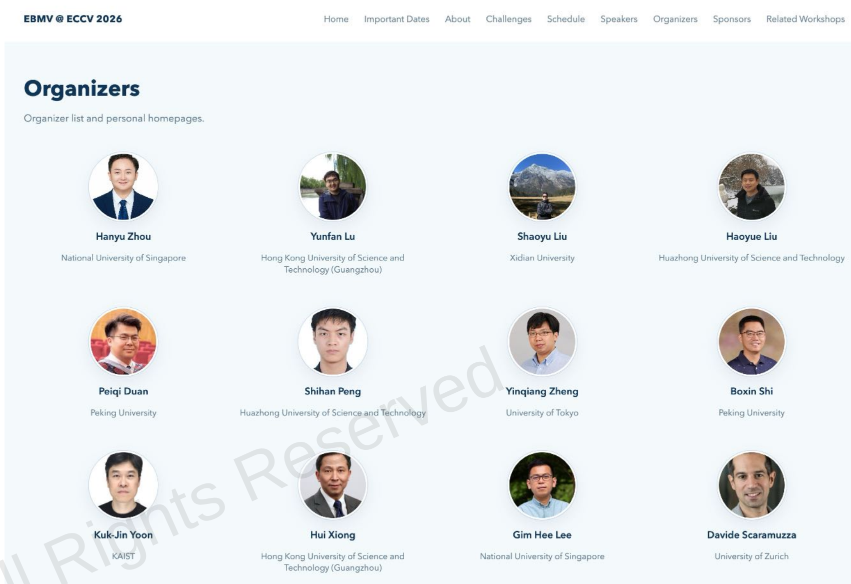
- Yunfan LU, Xiaogang Xu, Pengteng Li, Yusheng Wang, Yi Cui, Huizai Yao, Hui Xiong; From Events to Enhancement: A Survey on Event-Based Imaging Technologies
- Yunfan LU, Yiqi Lin, Hao Wu, Yunhao Luo, Xu Zheng, Hui Xiong; Priors in Deep Image Restoration and Enhancement: A Survey

Vision Foundation Model

- Pengteng Li, Yunfan LU, Pinhao Song, Weiyu Guo, Huizai Yao, F Richard Yu, Hui Xiong; DeblurSplat: SfM-free 3D Gaussian Splatting with Event Camera for Robust Deblurring, IEEE T-MM. CCF-A
- Pengteng Li, Yunfan LU, Pinghao Song, Wuyang Li, Huizai Yao, Hui Xiong; Eventvl: Understand event streams via multimodal large language model. Under Review
- Xiaogang Xu, Jian Wang, Yunfan LU, Ruihang Chu, Ruixing Wang, Jiafei Wu, Bei Yu, Liang Lin;Boosting Fidelity for Pre-Trained-Diffusion-Based Low-Light Image Enhancement via Condition Refinement. Under Review

Public Services

- Reviewer for multiple international journals and conferences
- Organizer of event-based vision workshops (program coordination, invited speakers, challenge tracks)
- Contributed to community building and academic exchange in event-based vision




IEEE TRANSACTIONS ON
PATTERN ANALYSIS AND MACHINE INTELLIGENCE
2024 Reviewer Certificate
Yunfan Lu
This certificate is awarded in recognition of the recipient's efforts as a peer reviewer.
IEEE Transactions on Pattern Analysis and Machine Intelligence
2024 Peer Reviewer
Kyoung Mu Lee
Editor-in-Chief

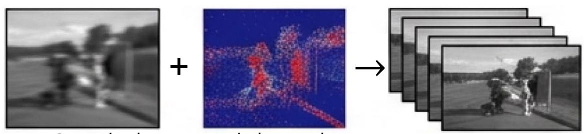

IEEE TRANSACTIONS ON
PATTERN ANALYSIS AND MACHINE INTELLIGENCE
2025 Reviewer Certificate
Yunfan Lu
This certificate is awarded in recognition of the recipient's efforts as a peer reviewer.
IEEE Transactions on Pattern Analysis and Machine Intelligence
2025 Peer Reviewer
Kyoung Mu Lee
Editor-in-Chief

https://yunfanlu.github.io/

Overview

Temporal Resconstruction

Unified RS correction, deblurring, and interpolation



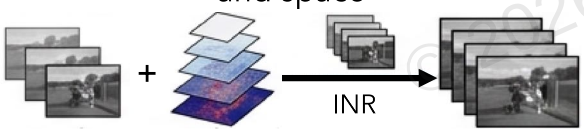
- Coupled temporal degradations
- One unified space-time representation
- Recover sharp global-shutter video

Time / Space

Event Sensor

Spatial-Temporal Reconstruction

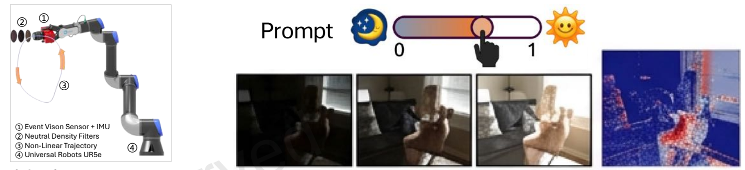
Continuous reconstruction across time and space



- Sparse frames + dense events
- Regional short-term motion
- Holistic long-term continuity

Dynamic Range Recovery

From fixed enhancement to controllable reconstruction

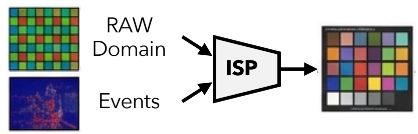


- Broad-light-range representation
- Brightness-controllable output
- Robust under complex illumination

Dynamic Range / RAW / Color / Sensor Modeling

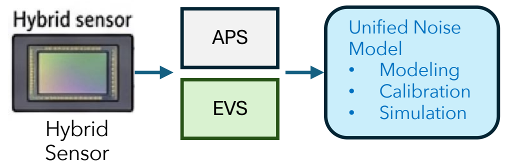
RAW / Color

Event-guided ISP beyond the RGB domain



- Events / RAW measurements
- More reliable color reconstruction
- Sensor-near imaging pipeline

Sensor Modeling



- Physics-informed statistical modeling
- Real-data calibration
- Simulation for downstream tasks

Event cameras enable a new imaging pipeline from sensing to reconstruction

Part II. Representative Works

From Image Reconstruction to Sensor Modeling

Five Representative Works

1. **Temporal Reconstruction**: UniINR: Event-Guided Unified Rolling Shutter Correction, Deblurring, and Interpolation. ECCV 2024
2. **Spatial-Temporal Reconstruction**: Continuous Space-Time Video Super-Resolution via Event Camera, IJCV 2025
3. **Dynamic Range Recovery**: SEE: See Everything Every Time, IJCV 2025
4. **RAW and Color Imaging**: RGB-Event ISP: Dataset, Benchmark, and Direction, ICLR 2025
5. **Sensor Modeling and Simulation**: Hybrid Event-Frame Sensors: Modeling, Calibration, and Simulation

Temporal Reconstruction – *Background*

Key observation: temporal degradation is inherently coupled in frame-based cameras.

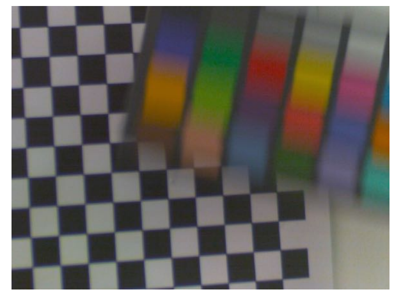
Under fast motion, the frame-based camera suffers from motion blur, rolling-shutter distortion, and missing temporal samples at the same time.

However, most existing methods treat these degradations separately, rather than solving them in a unified framework.

(a) Motion Blur

(b) Rolling Shutter Distortion

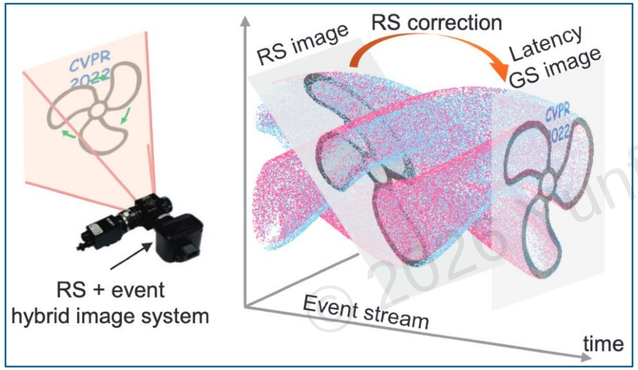
(c) Temporal Under Sampling



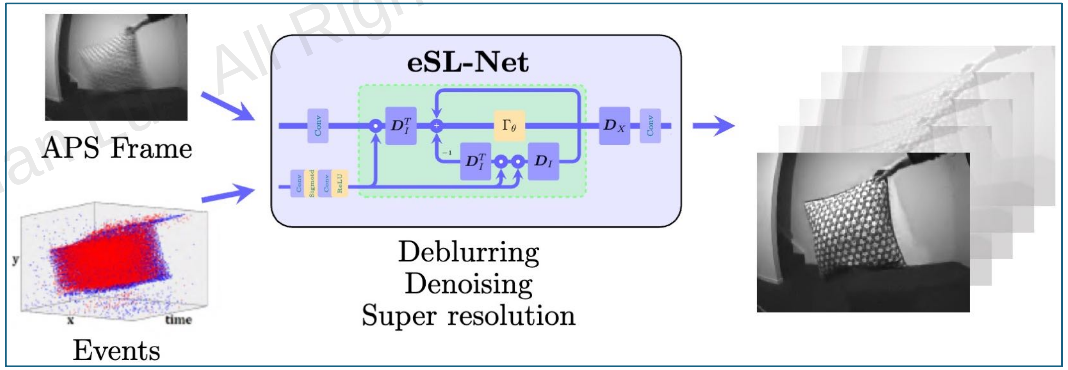
(d) They co-occur in real captures

Temporal Reconstruction – *Related Works*

- Most existing methods are staged, rather than jointly modeling coupled temporal degradations.
  - SelfUnroll[1], EvUnRoll[2], EvShutter[3] consider RS and blur, but often rely on staged designs, which increase model size and cumulative errors.
  - Most VFI methods, TimeLens[4], eSL-Net[5], DeblurSR[6], assume GS-style inputs or ignore RS distortion, so performance drops on spatially distorted RS blur frames.
- **Can we unify RS correction, deblurring, and interpolation in one representation?**



(a) RS Correction + Deblurring

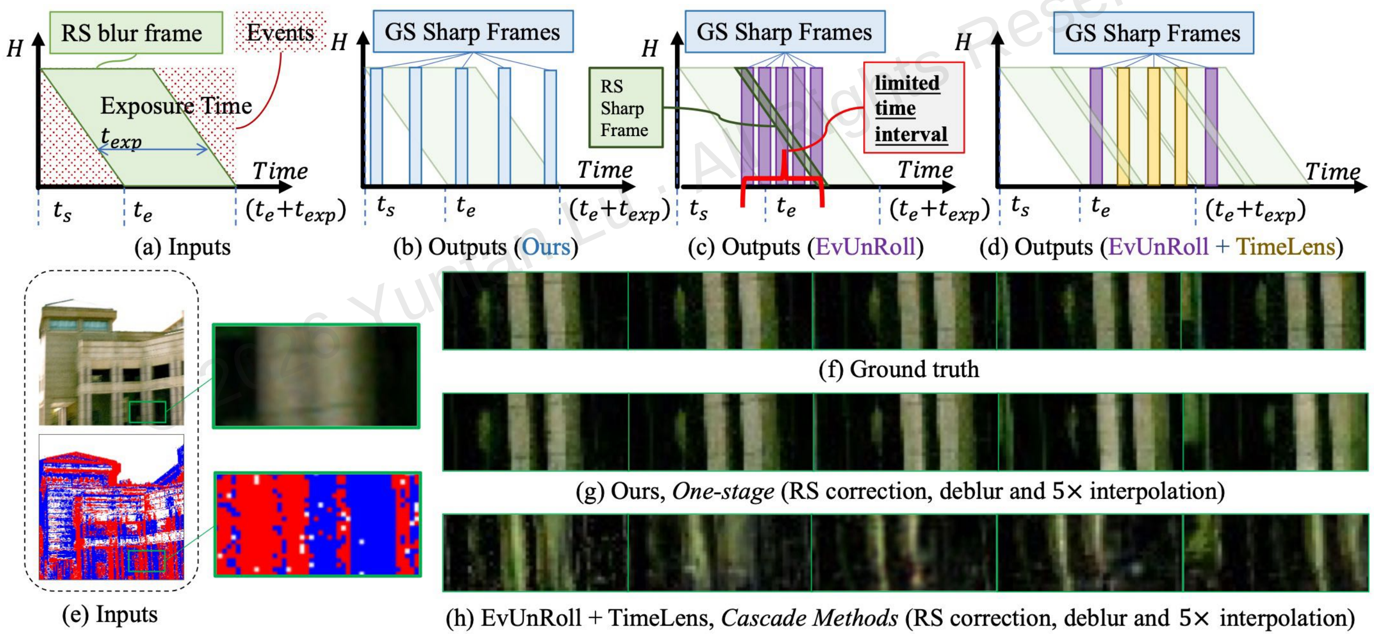


(b) Deblurring + Video Frame Interpolation

[1] Self-supervised scene dynamic recovery from rolling shutter images and events. IJCV 2024.
[2] Evunroll: Neuromorphic events based rolling shutter image correction. CVPR 2022.
[3] Evshutter: Transforming events for unconstrained rolling shutter correction. CVPR 2023
[4] Time lens: Event-based video frame interpolation. CVPR 2021
[5] Event enhanced high-quality image recovery. ECCV 2020
[6] Deblursr: event-based motion deblurring under the spiking representation. AAAI 2024

Temporal Reconstruction – *Key Insight*

- Key insight: Use one intermediate representation to unify multiple temporal degradations.
- We formulate reconstruction as a Implicit Neural Representation, $F(x, t, \theta)$: given pixel position x and time t, directly predict the sharp intensity or color.
- **The next question is: how do we build this unified INR from an RS blur frame and events?**



https://yunfanlu.github.io/

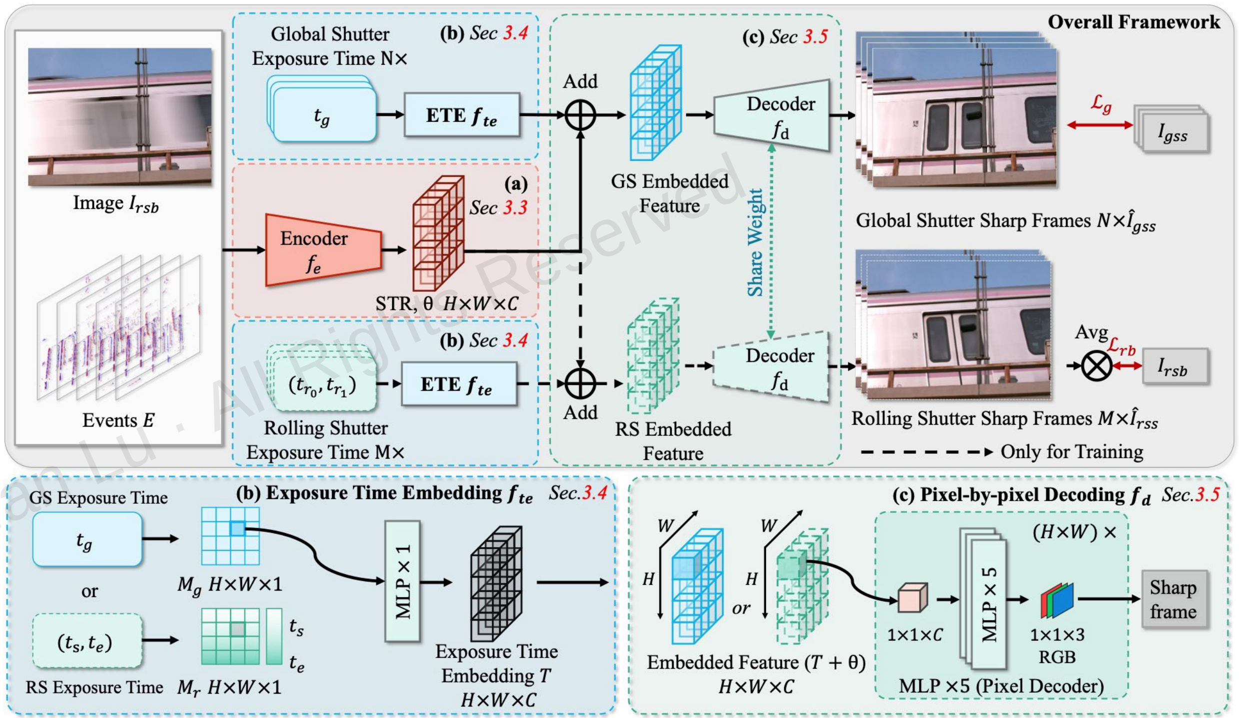
Temporal Reconstruction - *Method*

UniINR: a Unified Space-Time Representation

We implement the unified INR with three components:

- Spatial-Temporal Encoder
  - $\theta = f_e(I, E)$
- Exposure-Time Encoder
  - $\boldsymbol{t} = f_{te}(t)$
- Pixel-wise Prediction Decoder
  - $i_{out} = F(x, \boldsymbol{t}, \boldsymbol{\theta})$

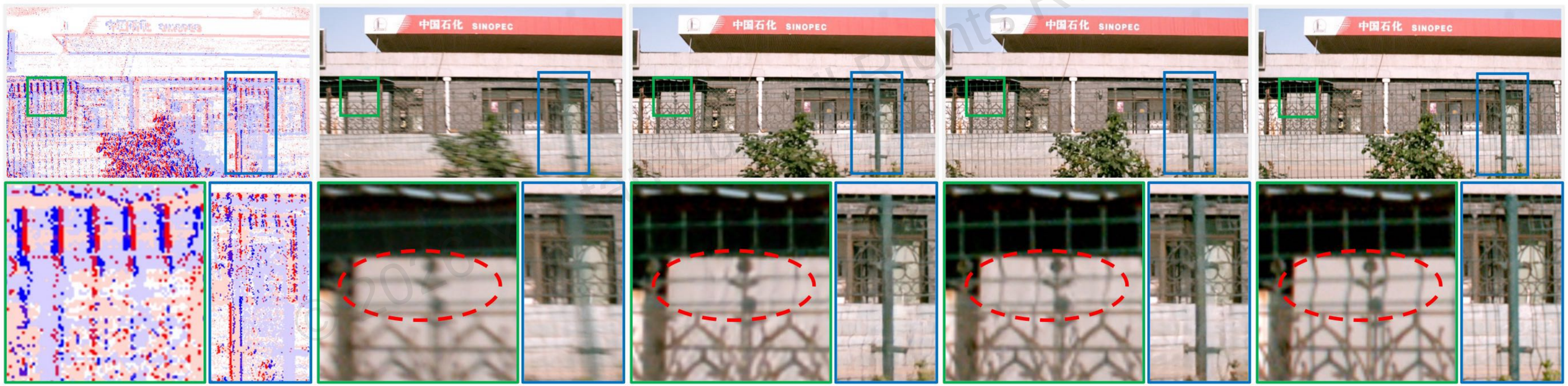
This design allows us to recover sharp frames at **arbitrary** timestamps with **low** complexity.



https://yunfanlu.github.io/

Temporal Reconstruction - *Experiments*

- Key result 1: One-stage unified reconstruction produces cleaner sharp videos than staged baselines.
  - <u>Better joint recovery</u>: One-pass recovery of RS, blur, and interpolation.
  - <u>Better efficiency-quality trade-off</u>: Only 0.38M parameters, yet better results.

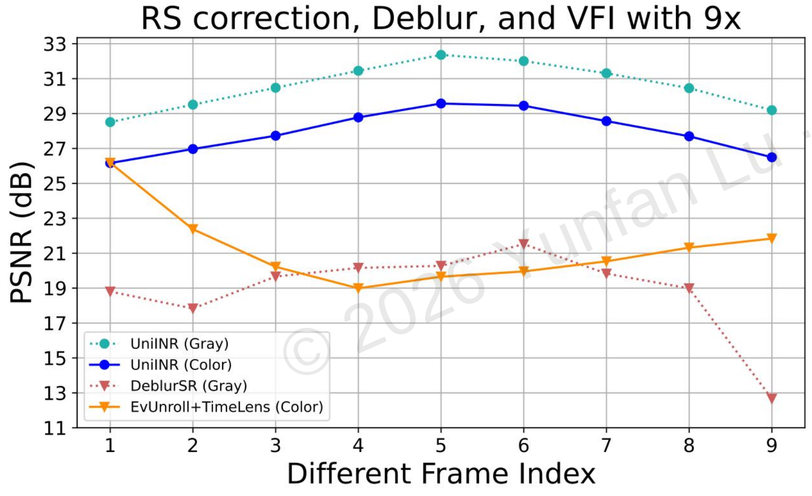


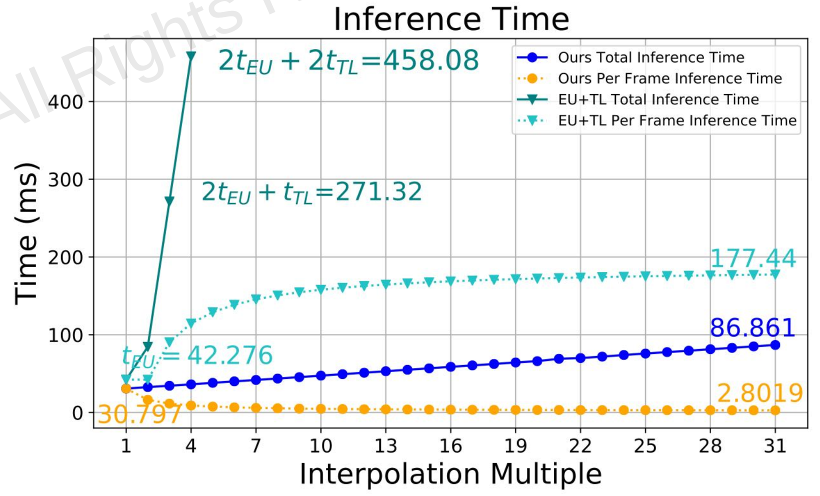
(a) Events (b) Input (c) EvUnRoll (d) Ours (e) GT

Please refer to the paper for more experiments.

Temporal Reconstruction - *Experiments*

- Key result 2: The unified formulation is also simple and efficient.
  - More stable reconstruction under large interpolation factors.
  - Fast per-frame decoding: only 2.8 ms/frame at 31× interpolation.





Temporal Reconstruction - *Summary*

1. Clearer formulation + physics-aware modeling = simpler and more efficient image reconstruction.

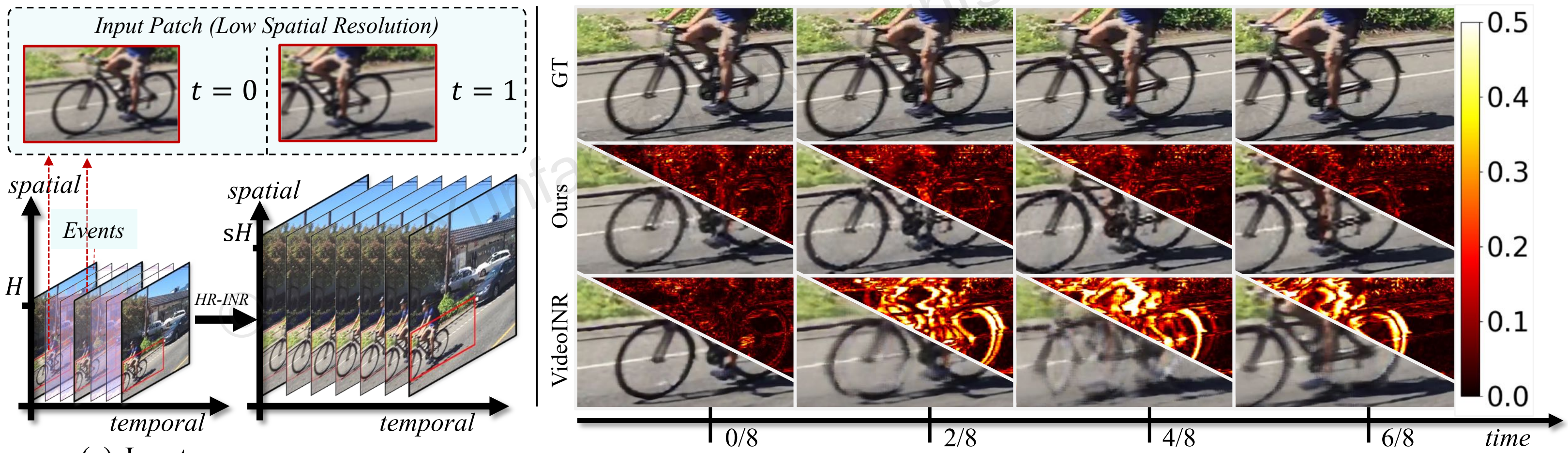
Five Representative Works

1. **Temporal Reconstruction**: UniINR: Event-Guided Unified Rolling Shutter Correction, Deblurring, and Interpolation. ECCV 2024
2. **Spatial-Temporal Reconstruction**: Continuous Space-Time Video Super-Resolution via Event Camera, IJCV 2025
3. **Dynamic Range Recovery**: SEE: See Everything Every Time, IJCV 2025
4. **RAW and Color Imaging**: RGB-Event ISP: Dataset, Benchmark, and Direction, ICLR 2025
5. **Sensor Modeling and Simulation**: Hybrid Event-Frame Sensors: Modeling, Calibration, and Simulation

https://yunfanlu.github.io/

Spatial-Temporal Rec. – *Background*

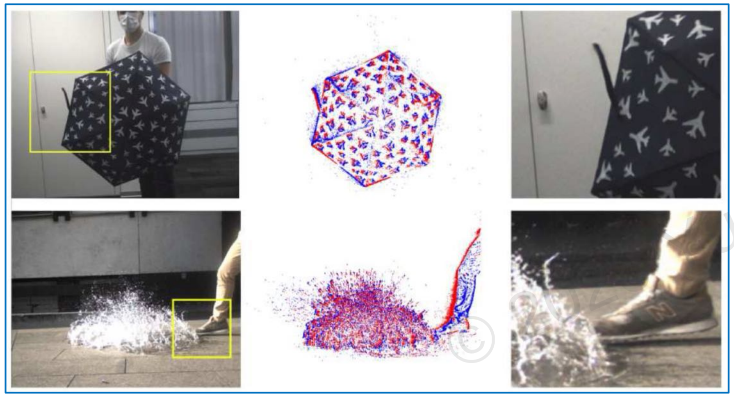
- Key observation: Real-world motion is continuous in both space and time, but cameras only sample it at low frame rates and fixed resolutions.
- Opportunity: Event cameras provide dense motion cues between sparse frames, and have shown value for both VFI and VSR.
- **Can events enable continuous space-time video super-resolution (C-STVSR)?**



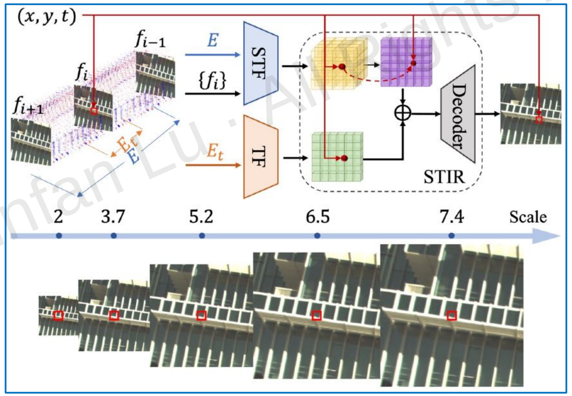
(a) Inputs *multiple frames* **(b)** Outputs **(c)** Visual Results, 4× SR and 8× VFI with Difference Map

Spatial-Temporal Rec. – *Related Works*

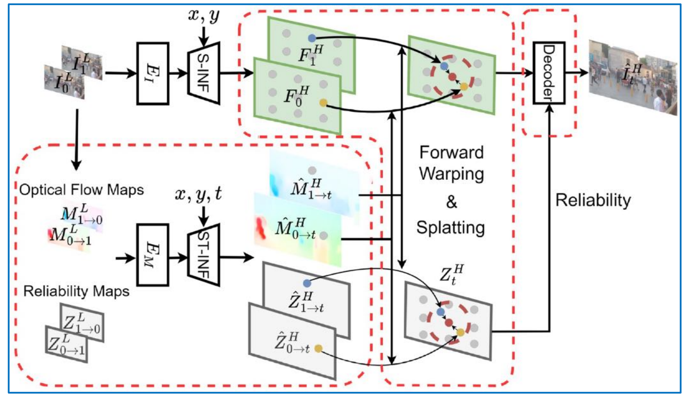
- Event-based VFI [1,2]: good at modeling nonlinear motion in time.
- Event-based VSR [3]: converts high-frequency temporal cues into high-frequency spatial details.
- Frame-based C-STVSR [4,5]: supports continuous decoding, but motion estimation is ill-posed.
- **Core takeaway: The key to event-enhanced C-STVSR is motion modeling.**



(a) Event-based Video Frame Interpolation [1,2]



(b) Event-based Video SR [3]



(c) Frame-based CSTVSR [4,5]

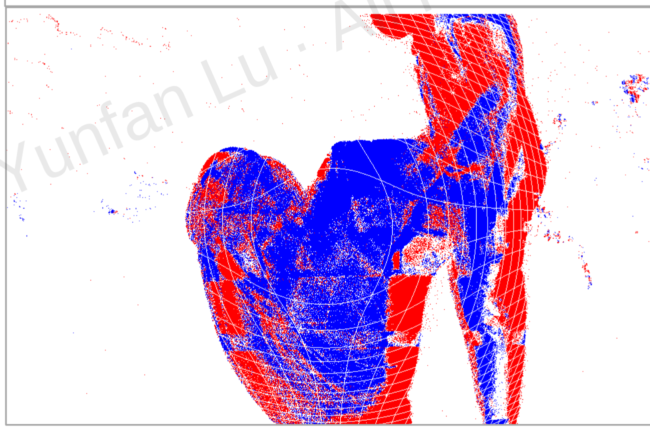
[1] Time lens++: Event-based frame interpolation with parametric non-linear flow and multi-scale fusion. CVPR 2022
[2] Event-based video frame interpolation with cross-modal asymmetric bidirectional motion fields. CVPR 2023
[3] Learning spatial-temporal implicit neural representations for event-guided video super-resolution. CVPR 2023
[4] Videoinr: Learning video implicit neural representation for continuous space-time super-resolution. CVPR 2022
[5] MoTIF: Learning motion trajectories with local implicit neural functions for continuous space-time video super-resolution. CVPR 2023

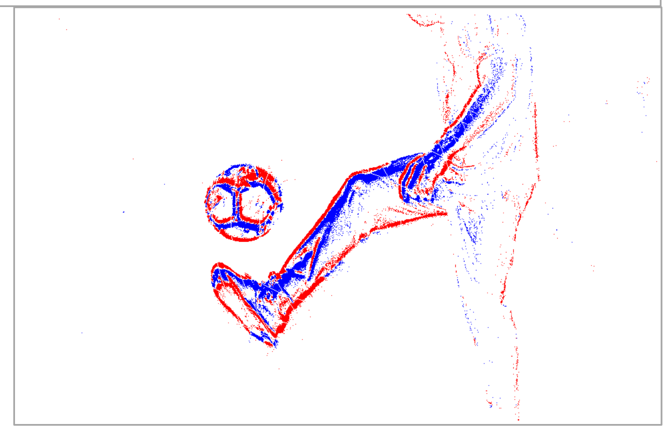
Spatial-Temporal Rec. – *Key Insight*

- Key insight: Motion modeling at two complementary scales: :
  - Holistic long-term continuity → multi-frame input
  - Regional short-term motion → short-term event representation
- We therefore design multi-frame event-frame fusion for long-term continuity, and fine-grained event temporal modeling for rapid local motion.

Events recorded almost continuous movement ~ *10,000 fps.*

(a) Sparse RGB frames miss motion states


(b) Holistic motion continuity

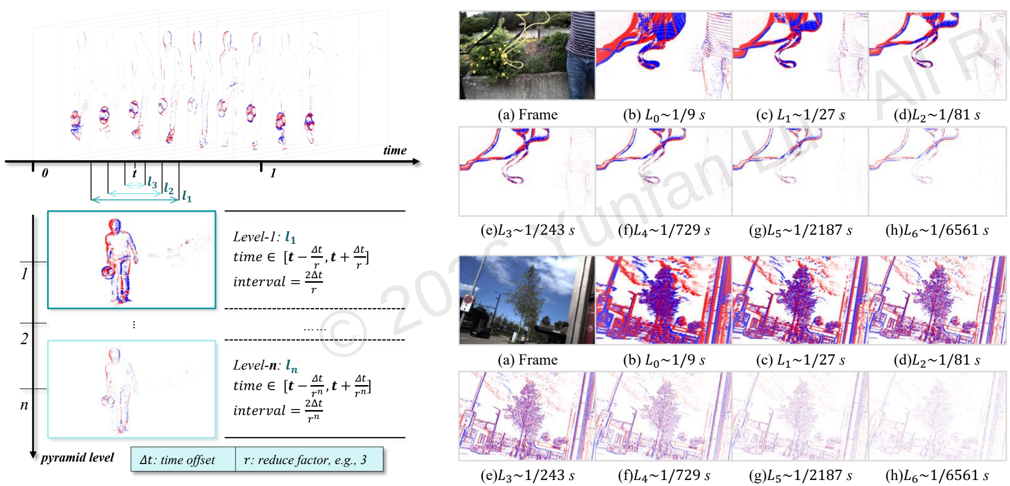

(c) Regional rapid motion

Spatial-Temporal Rec. – *Method*

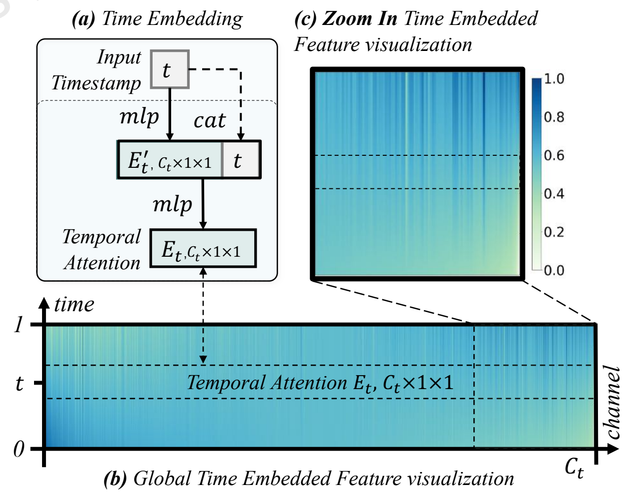
Regional short-term motion

- TPR: a pyramidal event representation for fine-grained short-term motion modeling.
- The finest temporal resolution reaches *1/6561s*.
- Time embedding injects timestamp information to avoid ambiguity in continuous decoding.

- Next, we move from short-term motion to holistic long-term continuity.



(a) Event Temporal Pyramid Representation

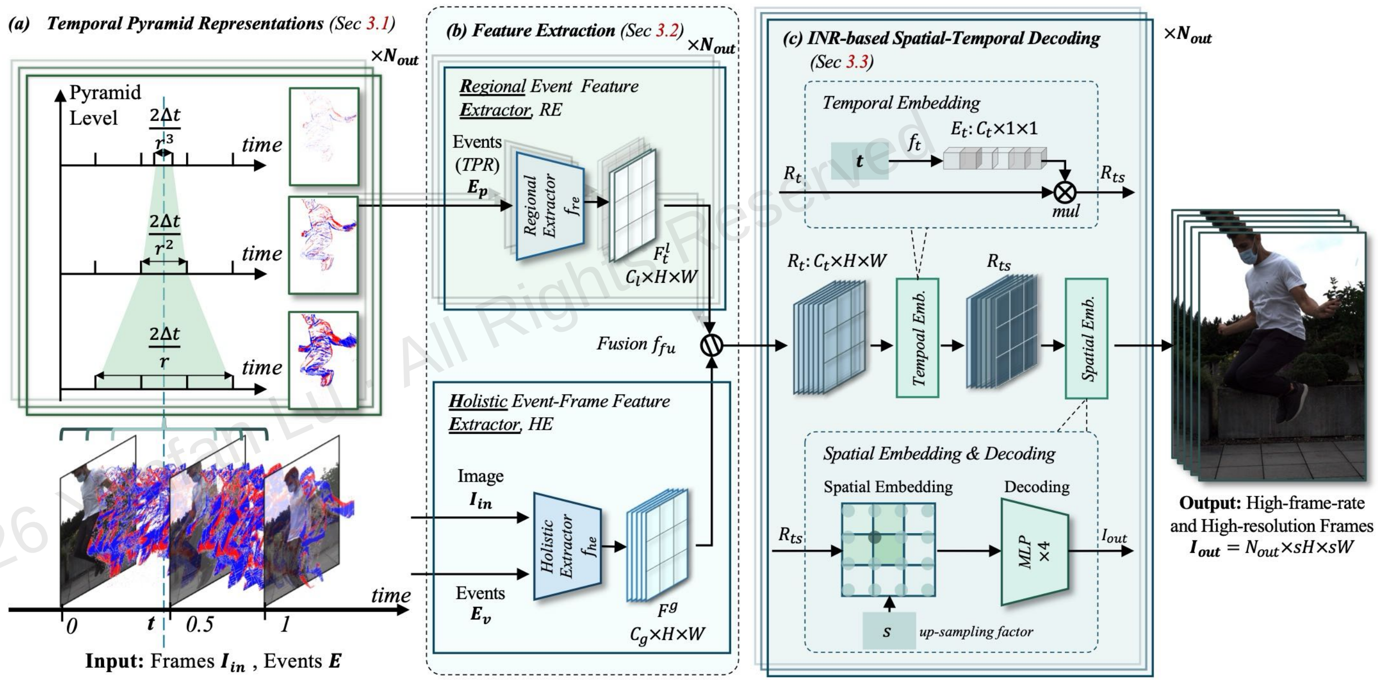


(b) Exposure Time Embedding

Spatial-Temporal Rec. – *Method*

Holistic long-term continuity

- Multi-frame input: move from two frames to multiple frames.
- HE fuses frames and events for long-term continuity.
- Temporal embedding selects the right long-term context for each timestamp.



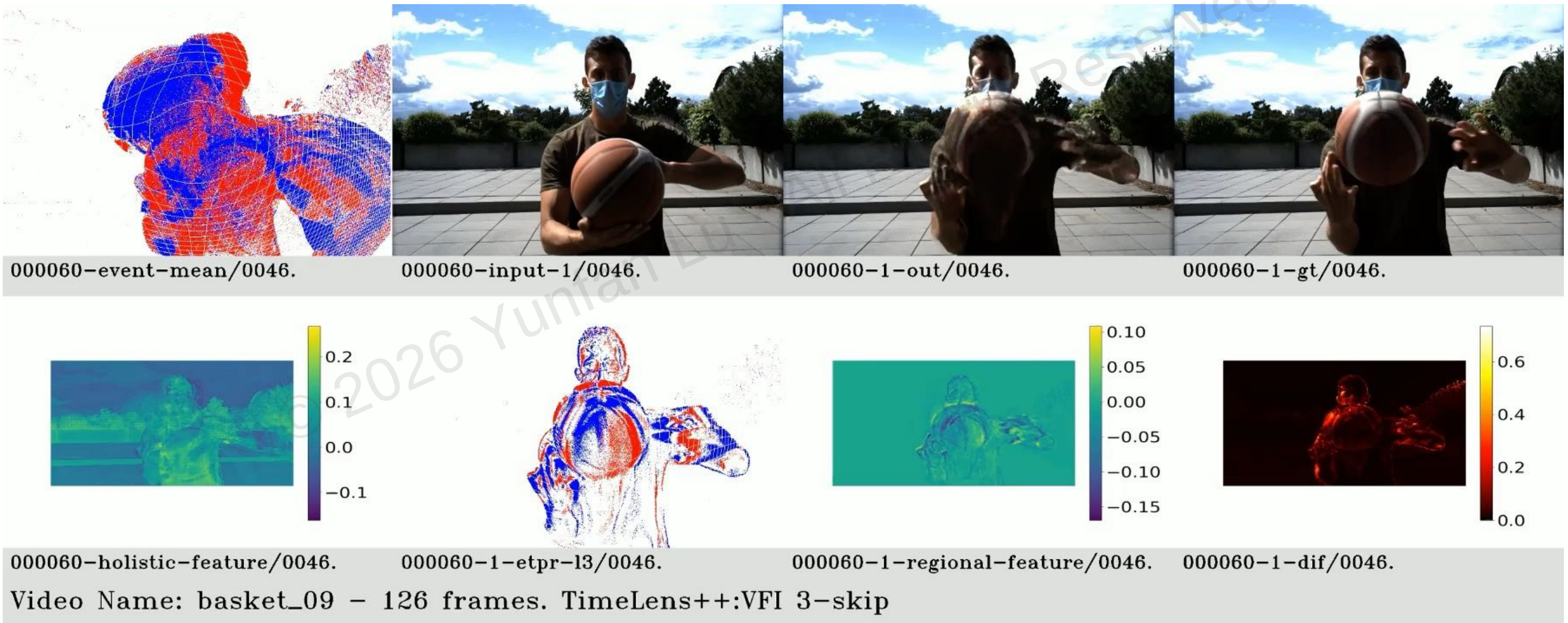
Spatial-Temporal Rec. – *Experiments*

- Result 1: Long-term temporal modeling and Event TPR jointly capture nonlinear motion.

https://yunfanlu.github.io/

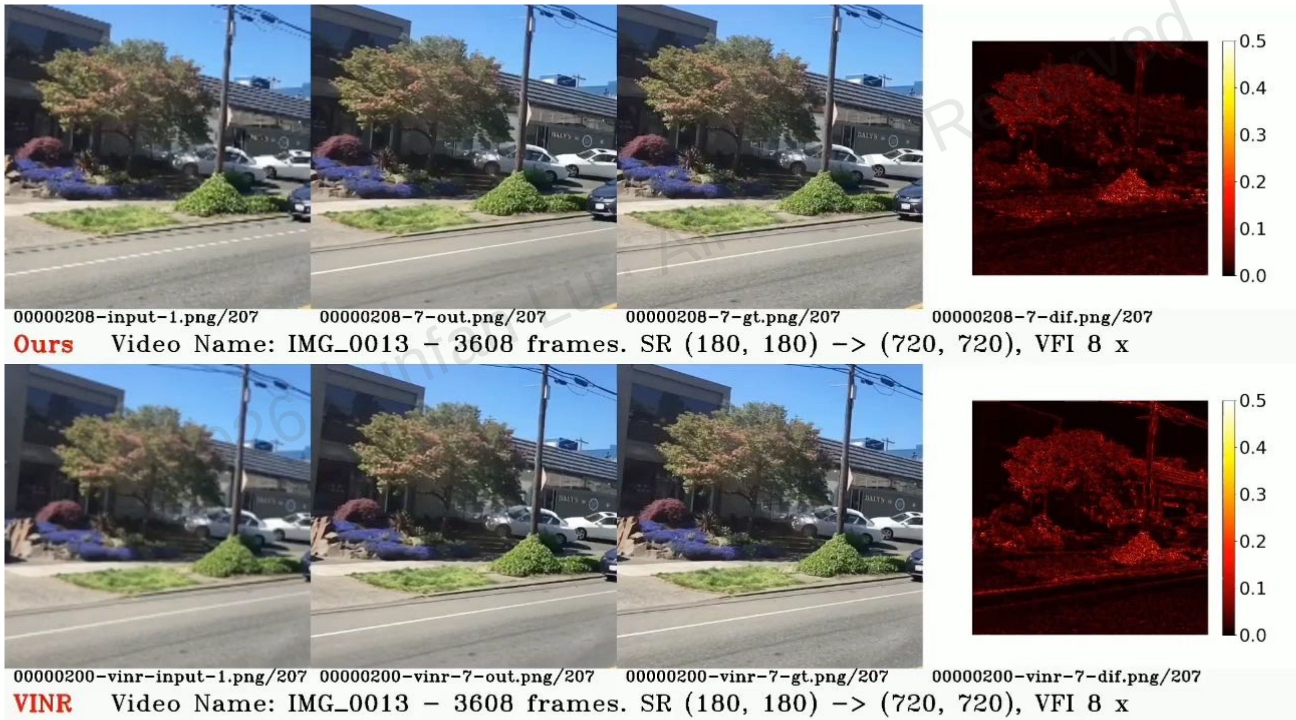
Spatial-Temporal Rec. – *Experiments*

- Result 1: Long-term temporal modeling and Event TPR jointly capture nonlinear motion.



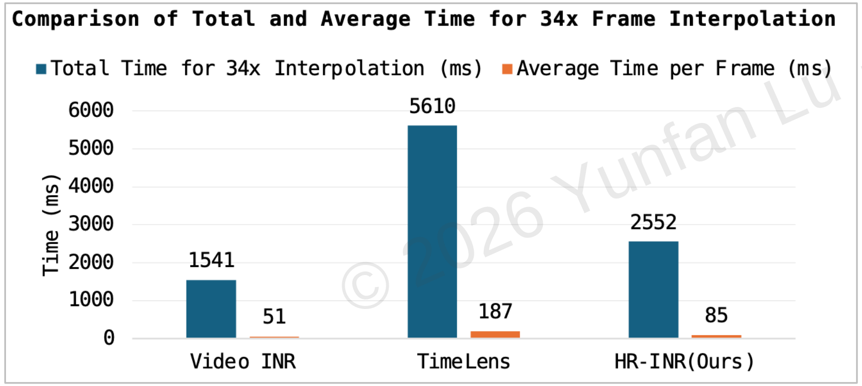
Spatial-Temporal Rec. – *Experiments*

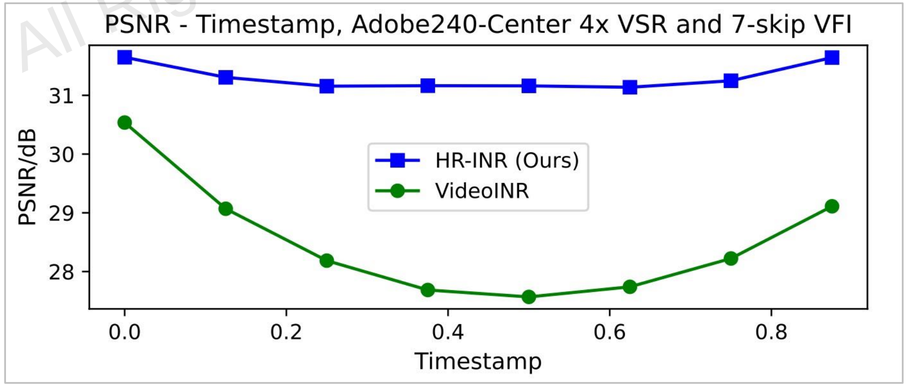
- Result 2: The joint design improves both single-task and multi-task performance.



Spatial-Temporal Rec. – *Experiments*

- Result 2: The joint design improves both single-task and multi-task performance.
  - Efficient under high interpolation factors
  - More stable quality across timestamps





Spatial-Temporal Rec. – *Summary*

1. Clearer formulation + physics-aware modeling = simpler and more efficient image reconstruction.
2. Better motion representation + more tailored architecture design = stronger continuous space-time reconstruction.

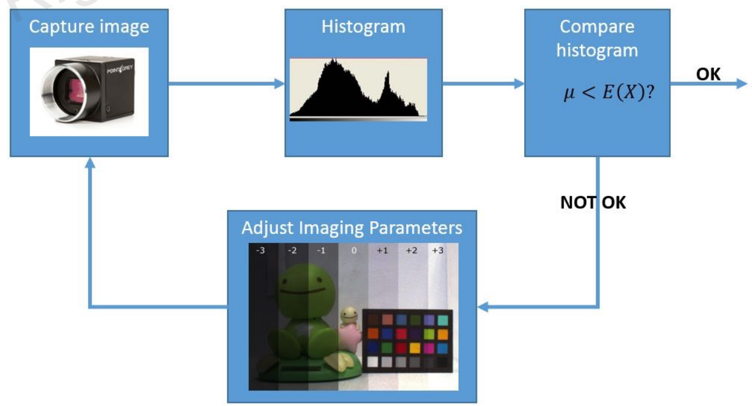
Five Representative Works

1. **Temporal Reconstruction**: UniINR: Event-Guided Unified Rolling Shutter Correction, Deblurring, and Interpolation. ECCV 2024
2. **Spatial-Temporal Reconstruction**: Continuous Space-Time Video Super-Resolution via Event Camera, IJCV 2025
3. **Dynamic Range Recovery**: SEE: See Everything Every Time, IJCV 2025
4. **RAW and Color Imaging**: RGB-Event ISP: Dataset, Benchmark, and Direction, ICLR 2025
5. **Sensor Modeling and Simulation**: Hybrid Event-Frame Sensors: Modeling, Calibration, and Simulation

Dynamic Range Rec. – *Background*

- Key observation: Real-world illumination (0.0003~100,000 *lux*) is far more complex. Auto-exposure (AE) of camera fails under extreme lighting conditions.
- Opportunity: Event cameras measure brightness changes with high dynamic range, making them less prone to overexposure or underexposure.
- **If events help under difficult lighting, what have existing event-based methods already solved – and what is still missing?**

(a) Real lighting conditions [1]

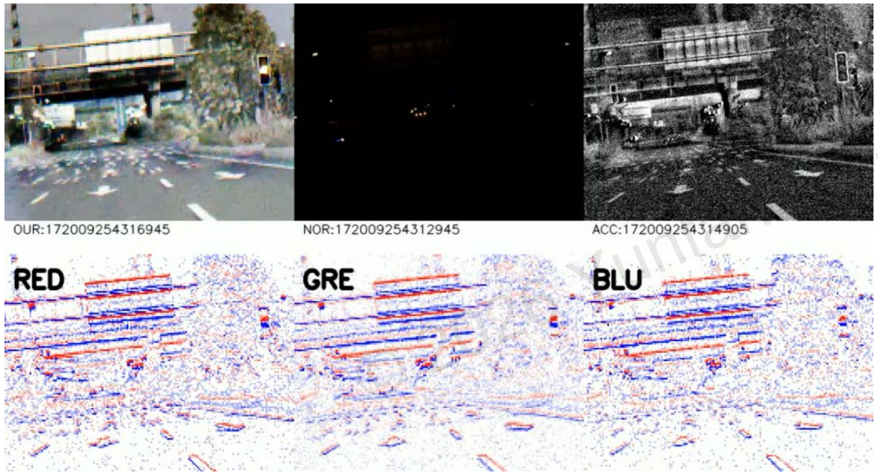


(b) Auto-exposure in traditional cameras

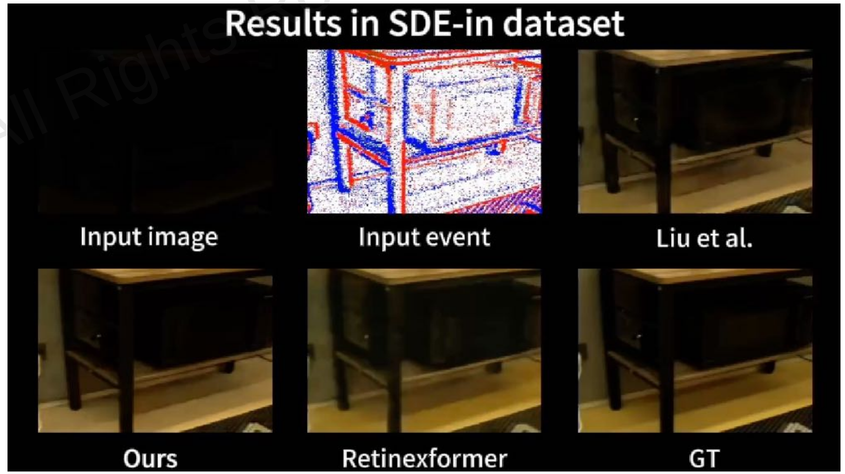
[1] Gehrig, Mathias, and Davide Scaramuzza. "Recurrent vision transformers for object detection with event cameras." CVPR (2023).

Dynamic Range Rec. – *Related Works*

- Prior works [1,2,3,4]: Events have been shown useful for low-light reconstruction and image enhancement.
- Limitation: Most methods still map one lighting condition to one fixed target brightness.
- This suggests a new direction: **instead of fixed enhancement, can we make brightness itself controllable?**



(a) Video reconstruction directly from events streaming

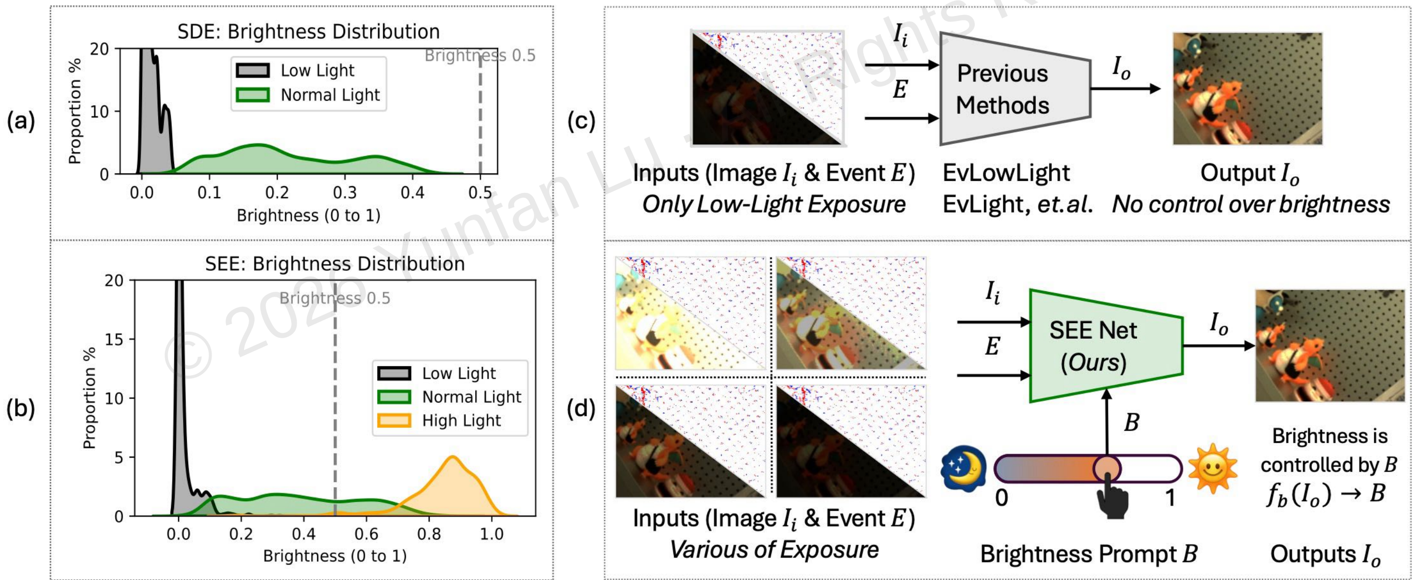


(b) Low light video enhancement from events and frames

[1] Towards robust event-guided low-light image enhancement: a large-scale real-world event-image dataset and novel approach. CVPR 2024
[2] Revisit Event Generation Model: Self-Supervised Learning of Event-to-Video Reconstruction with Implicit Neural Representations. ECCV 2024
[2] Coherent event guided low-light video enhancement. CVPR 2023
[3] EvLight++: Low-Light Video Enhancement With an Event Camera: A Large-Scale Real-World Dataset, Novel Method, and More. IEEE T-PMAI 2025

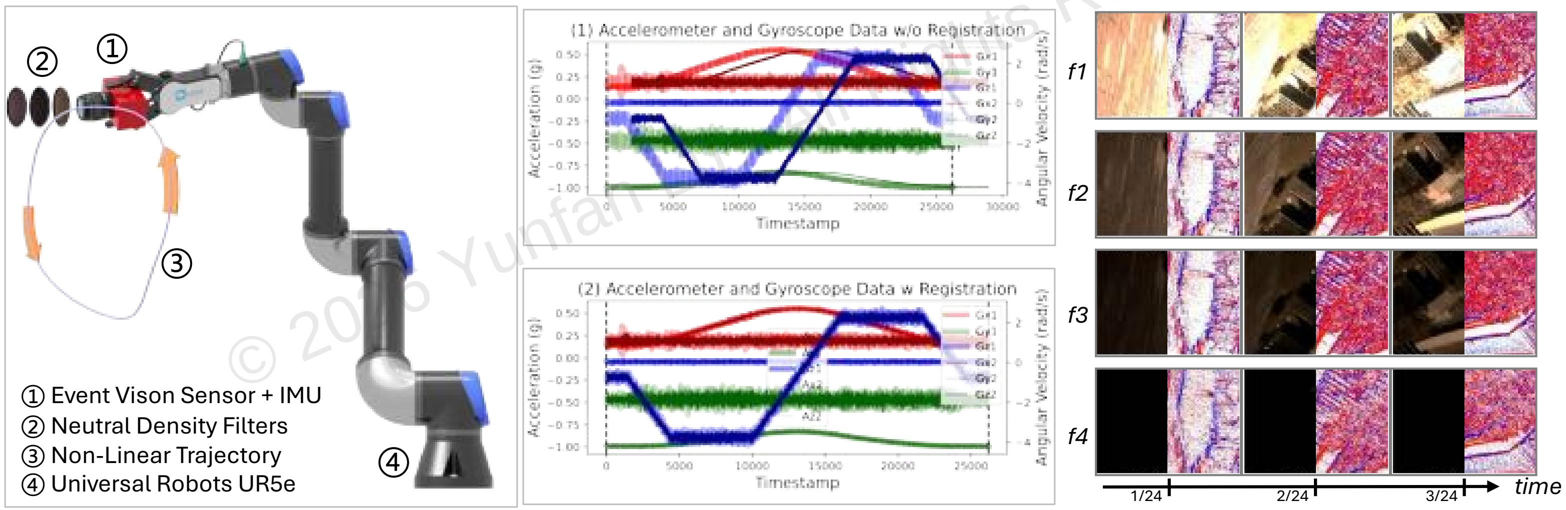
Dynamic Range Rec. – *Key Insight*

- Key insight: Move from fixed enhancement to controllable brightness adjustment.
- Learn a unified broad-light-range representation instead of a single target exposure.
- Control: Use a brightness prompt to specify the desired output brightness.
- **Data challenge: how do we build a multi-brightness aligned dataset for this task?**



Dynamic Range Rec. – *Dataset*

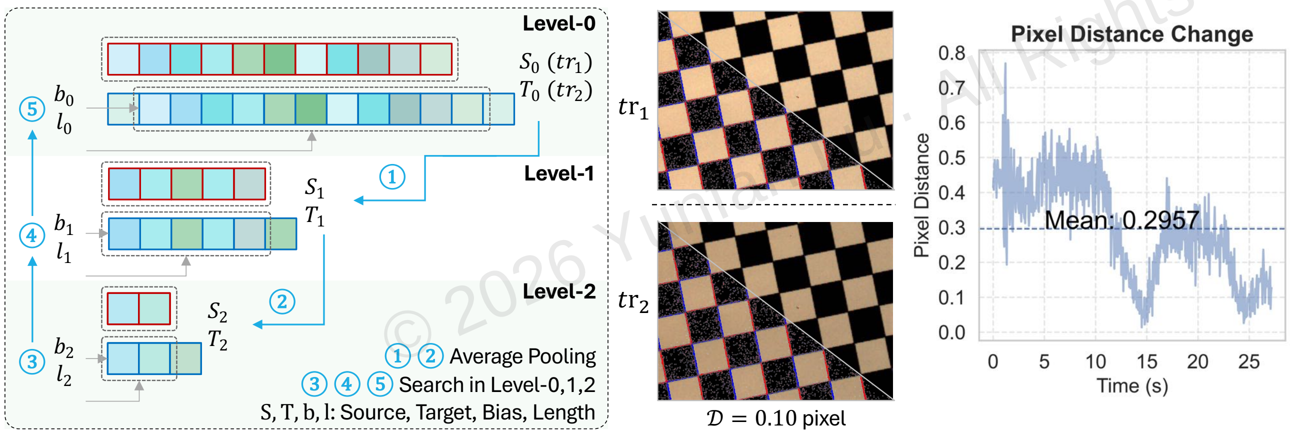
- Requirement: Broad-light-range learning needs frames captured under multiple brightness levels, with spatial-temporal alignment.
- Solution: We design an <u>IMU-based registration (*1000 Hz*)</u> pipeline for sub-pixel alignment across brightness settings.
- Outcome: This leads to SEE-600K, a large-scale dataset for broad illumination conditions.



(a) Data Collection Setup (b) IMU Data Registration (c) EVS Outputs with Different Filter

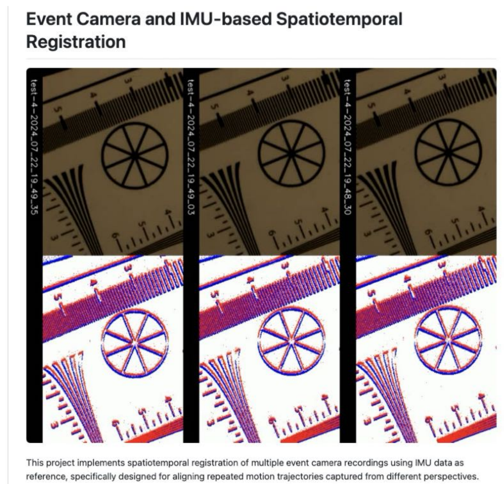
Dynamic Range Rec. – *Dataset*

- The IMU-based registration method achieves sub-pixel alignment, with about 0.29-pixel mean error between recordings.
- Implication: This makes it possible to build a reliable multi-brightness dataset for controllable brightness learning.
- **With reliable alignment in place, we can build SEE-600K for broad-light-range learning.**



(a) Registration Process (b) Two trajectories (c) Pixel Distance Change

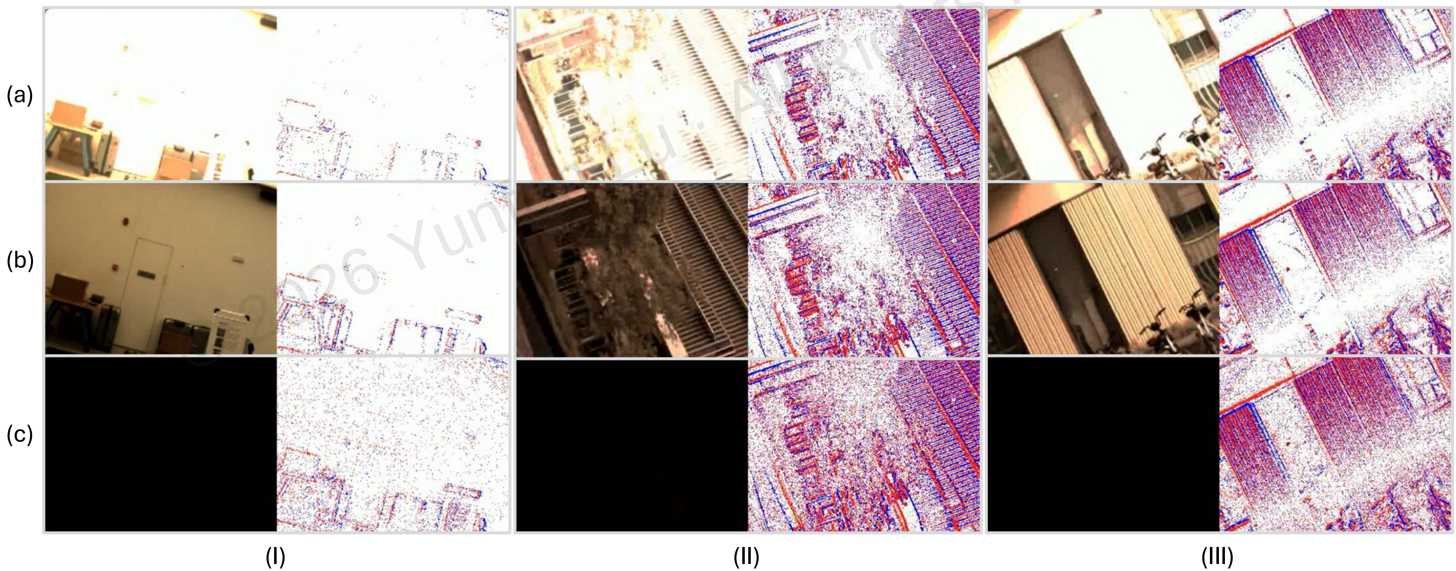
(a) Event Camera and IMU-based Spatiotemporal Registration



(b) IMU-Registration-Tool [2]

Dynamic Range Rec. – *Dataset*

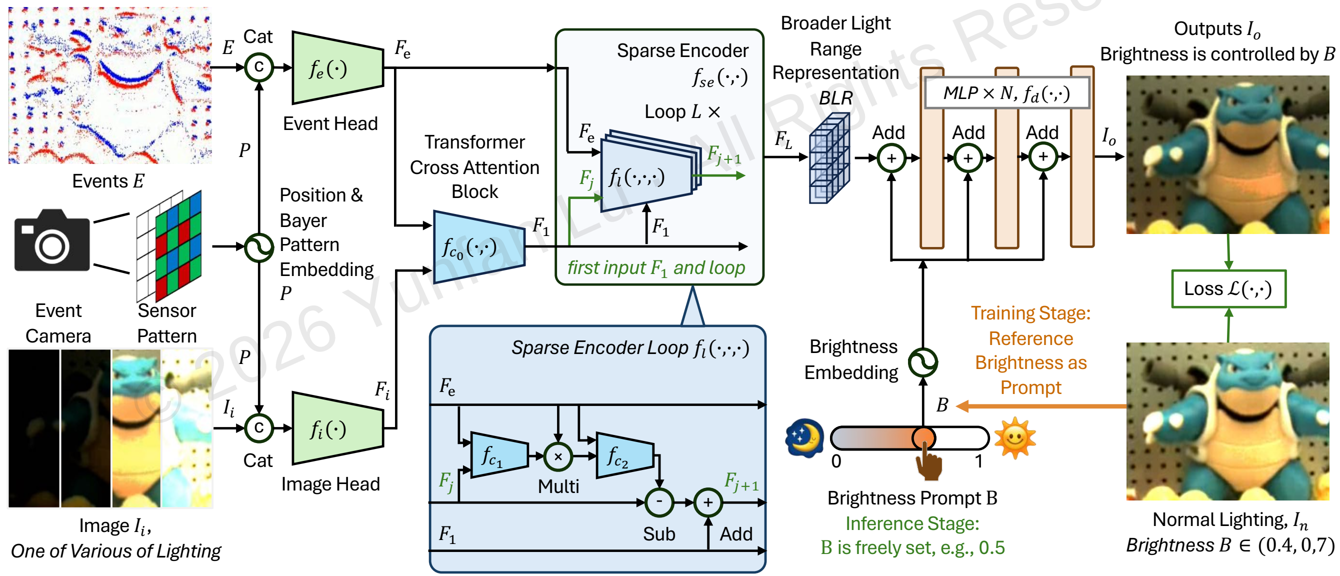
- SEE-600K: aligned data across low-, normal-, and high-light conditions.
- Same scene, different brightness levels, same spatial-temporal content.
- This enables supervised learning of broad-light-range imaging.
- **The next step is to design a model that learns this broad-light-range representation.**



https://yunfanlu.github.io/

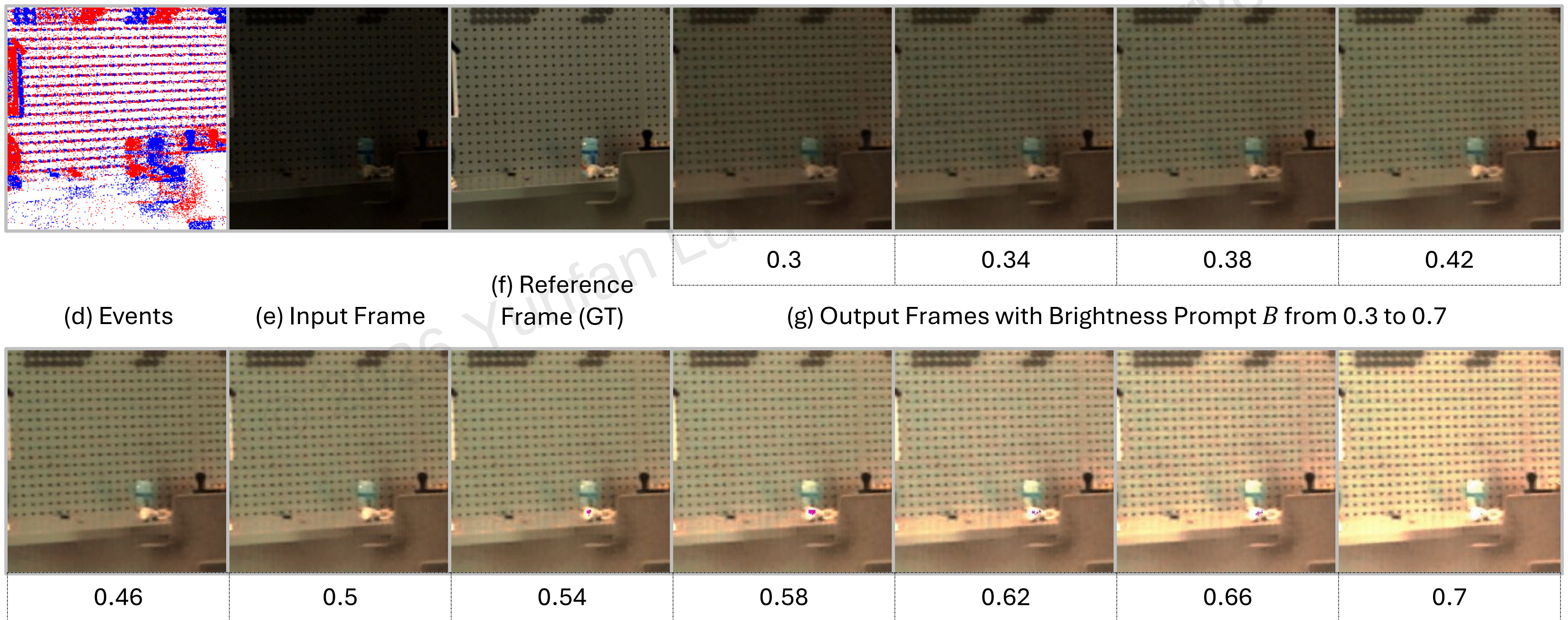
Dynamic Range Rec. – *Method*

- Key idea: From fixed enhancement to controllable brightness adjustment.
- Previous: $I_o = f(I_i, E)$. Ours: $I_o = f(I_i, E, B)$
- B specifies the target brightness.



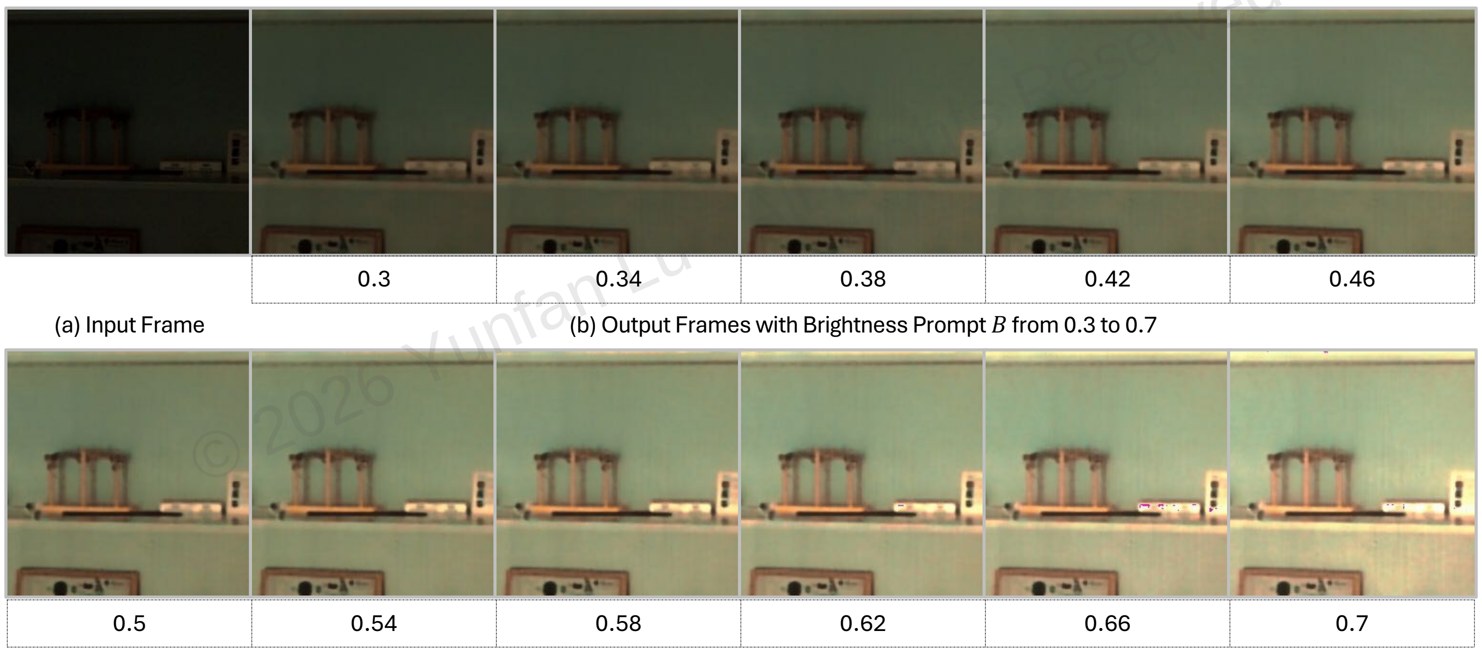
Dynamic Range Rec. – *Experiments*

- Key Result 1: SEE-Net enables smooth and controllable brightness adjustment.
- As the brightness prompt B varies from 0.3 to 0.7, the output transitions smoothly from dark to bright, demonstrating precise brightness control and broad-light-range modeling.



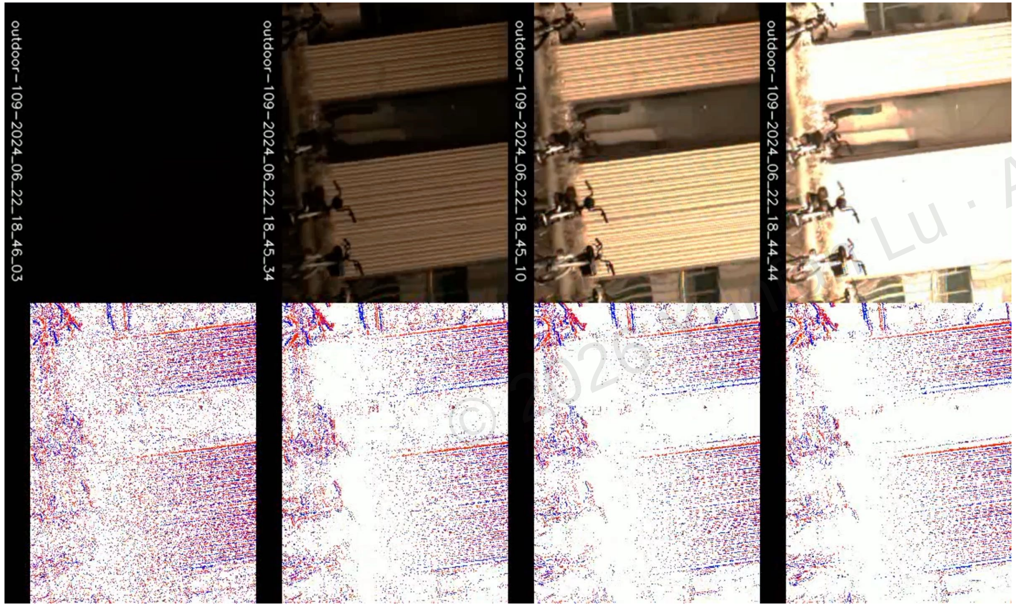
Dynamic Range Rec. – *Experiments*

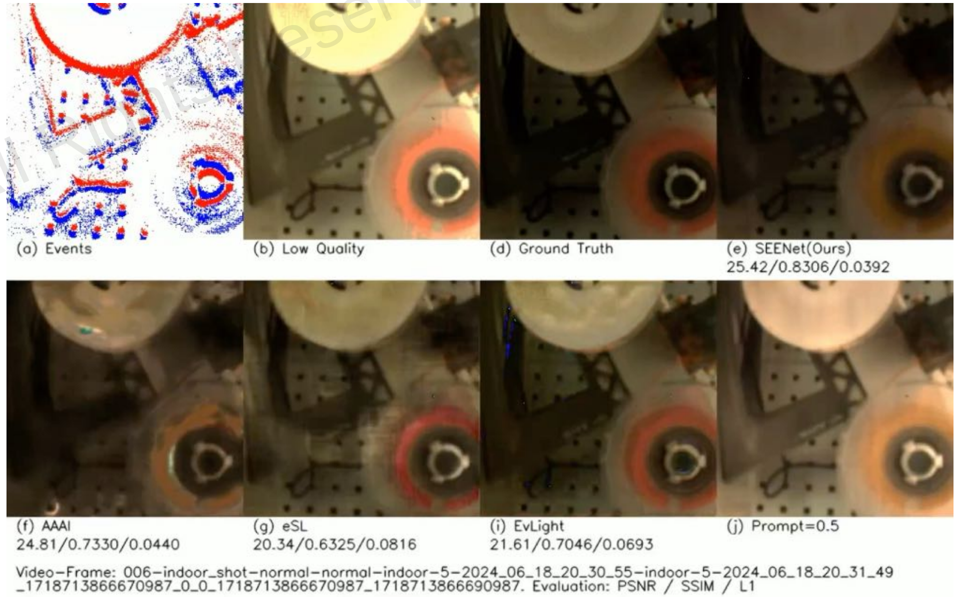
- Key Result 1: SEE-Net enables smooth and controllable brightness adjustment.
- As the brightness prompt B varies from 0.3 to 0.7, the output transitions smoothly from dark to bright, demonstrating precise brightness control and broad-light-range modeling.



Dynamic Range Rec. – *Experiments*

- Key Result 2: A controllable task definition leads to better reconstruction.
- By removing the ambiguity of direct low-light-to-normal mapping, SEE-Net achieves stronger results across diverse lighting conditions.





Dynamic Range Rec. – *Summary*

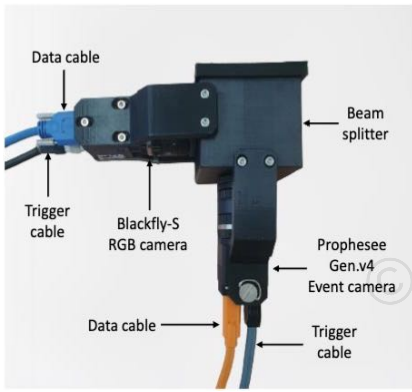
1. Clearer question formulation + physics-aware modeling = simpler and more efficient image reconstruction.
2. Better motion representation + more tailored architecture design = stronger continuous space-time reconstruction.
3. New dataset construction + new task formulation = a broader event-based imaging paradigm.

Five Representative Works

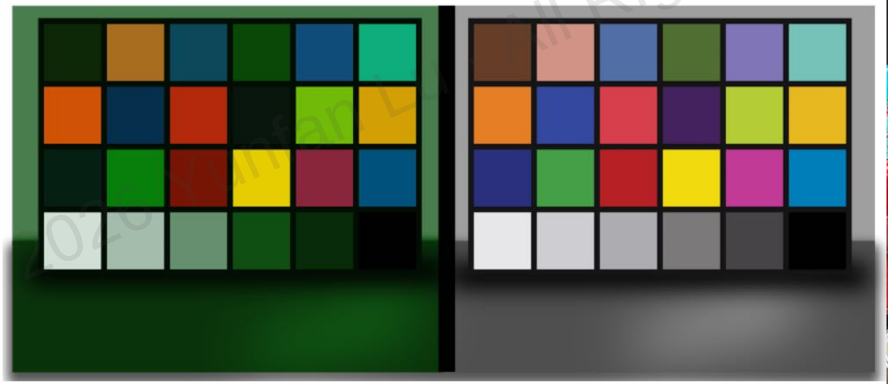
1. **Temporal Reconstruction**: UniINR: Event-Guided Unified Rolling Shutter Correction, Deblurring, and Interpolation. ECCV 2024
2. **Spatial-Temporal Reconstruction**: Continuous Space-Time Video Super-Resolution via Event Camera, IJCV 2025
3. **Dynamic Range Recovery**: SEE: See Everything Every Time, IJCV 2025
4. **RAW and Color Imaging**: RGB-Event ISP: Dataset, Benchmark, and Direction, ICLR 2025
5. **Sensor Modeling and Simulation**: Hybrid Event-Frame Sensors: Modeling, Calibration, and Simulation

RAW and Color - *Background*

- Key observation: Existing event-guided methods mainly operate in the RGB domain, rather than in the sensor-near RAW domain.
- Opportunity: Events preserve **richer** cues about temporal variation and dynamic illumination.
- Question: **Can events enhance the ISP pipeline, for example, by recovering more accurate and robust colors?**



Data acquisition devices

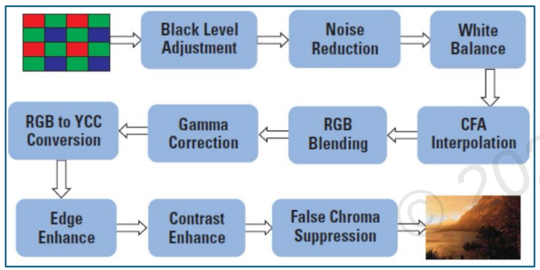


Images without and with an ISP

The real world is full of colors.

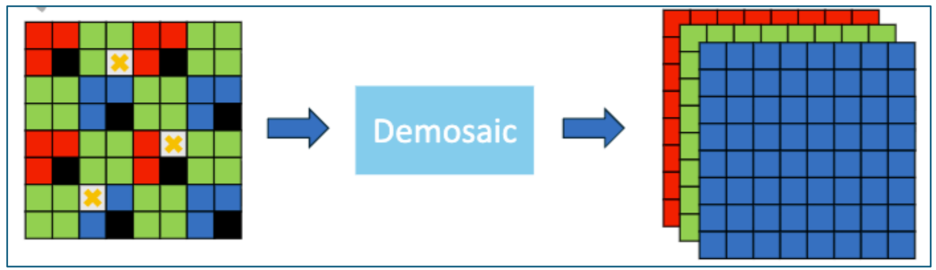
RAW and Color – *Related Works*

- Traditional ISP: handles multiple sub-tasks with carefully engineered modules and strong domain knowledge [1].
- Learning-based ISP: learns the pipeline end-to-end, requires large-scale paired datasets [2].
- Events in the RAW domain: remain largely underexplored, with only preliminary attempts on demosaicing, rather than the full ISP pipeline [3,4].
- **The next step is to build a dataset for event-guided RAW and color imaging.**



Traditional ISP

Learned ISP

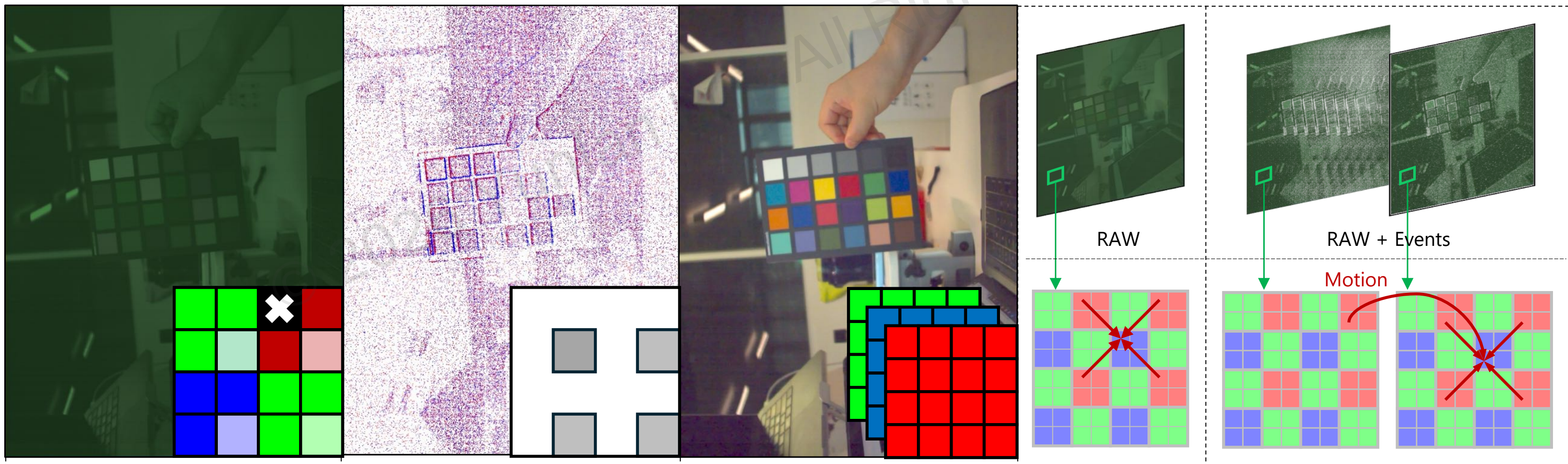


CVPR 2024 MIPI Workshop (Demosaic)

[1] Toward learning an end-to-end image processing pipeline. IEEE TIP 2018
[2] Replacing mobile camera isp with a single deep learning model. CVPR 2020
[3] Mipi 2024 challenge on demosaic for hybridevs camera: Methods and results. CVPR 2024
[4] Event Camera Demosaicing via Swin Transformer and Pixel-focus Loss. CVPRW 2024

RAW and Color – *Key Insights*

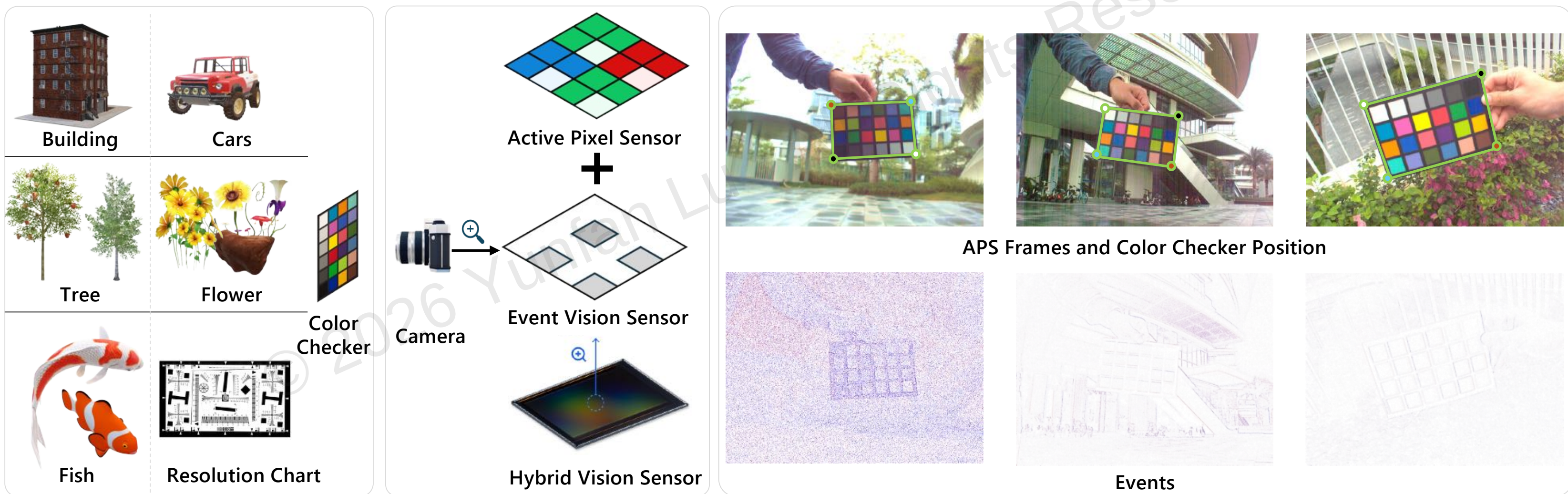
- Events provide complementary information beyond RAW.
- This makes event-guided ISP feasible for color recovery.
- The real bottleneck is dataset construction.
- **With such a dataset in place, we can finally ask whether events truly help the ISP pipeline.**



(a) RAW (b) Events (c) RGB Frame (d) ISP (e) Event-guided ISP

RAW and Color – *Dataset*

- A color checker is introduced as an explicit physical color reference.
- It guides controllable ISP toward more accurate color reconstruction.
- **We need an accurate color response pipeline.**



(a) Examples of Dataset Scenes (b) Camera (c) Examples of Our Dataset

RAW and Color – *Dataset*

Controllable ISP Pipeline

Black Level/FPN Removal
- Global blc subtraction + Row-wise fixed pattern

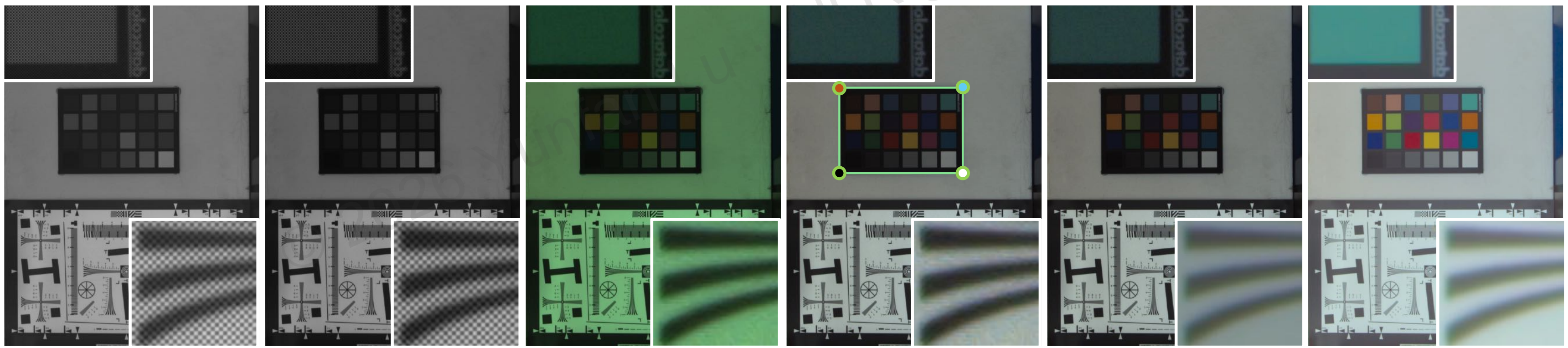
Demosaic
- Rainbow method → Preserves resolution
- Challenge: False colors in high-frequency areas

White Balance
- ColorChecker Patch #21 → Manual calibration

BM3D Denoising
- σ=50 setting for spatial noise reduction

Color Correction
- CIEDE00-optimized 3×3 CCM matrix

Gamma Encoding
- sRGB standard piecewise curve



(a) Quad-bayer pattern raw (b) Black pattern, Fixed-pattern (c) Demosacing (d) White Balancing (e) Denoise (f) Color space transform & Gamma

RAW and Color – *Dataset*

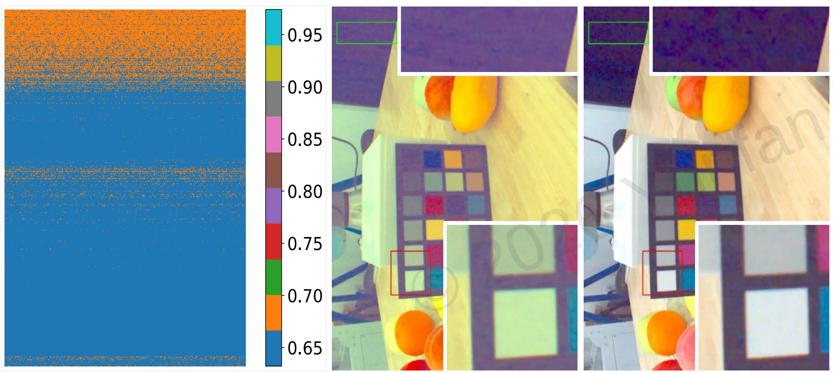
Controllable ISP Pipeline

- CIEDE00 (Human Perception):
Mean: 5.84 | Median: 5.07 → *Expert-grade accuracy*
Max Error: ~14 (sensor hardware limits)

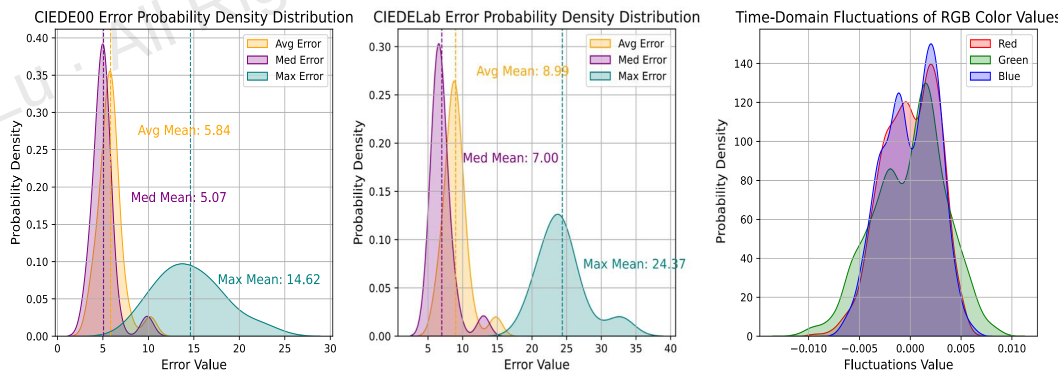
- CIELAB Δab:
Mean: 8.99 | Median: 7.00
Max Error: ~24 (photodiode layout impact)

- Temporal Consistency
140-Frame Video Test:
Color drift < 0.01 (all 24 patches)
90% fluctuations < 0.005 → Cinema-grade stability

This dataset has accurate and consistent colors to support benchmarking.



(a) Fixed Pattern Noise (FPN) (b) ISP w/o remove FPN (c) ISP w remove FPN

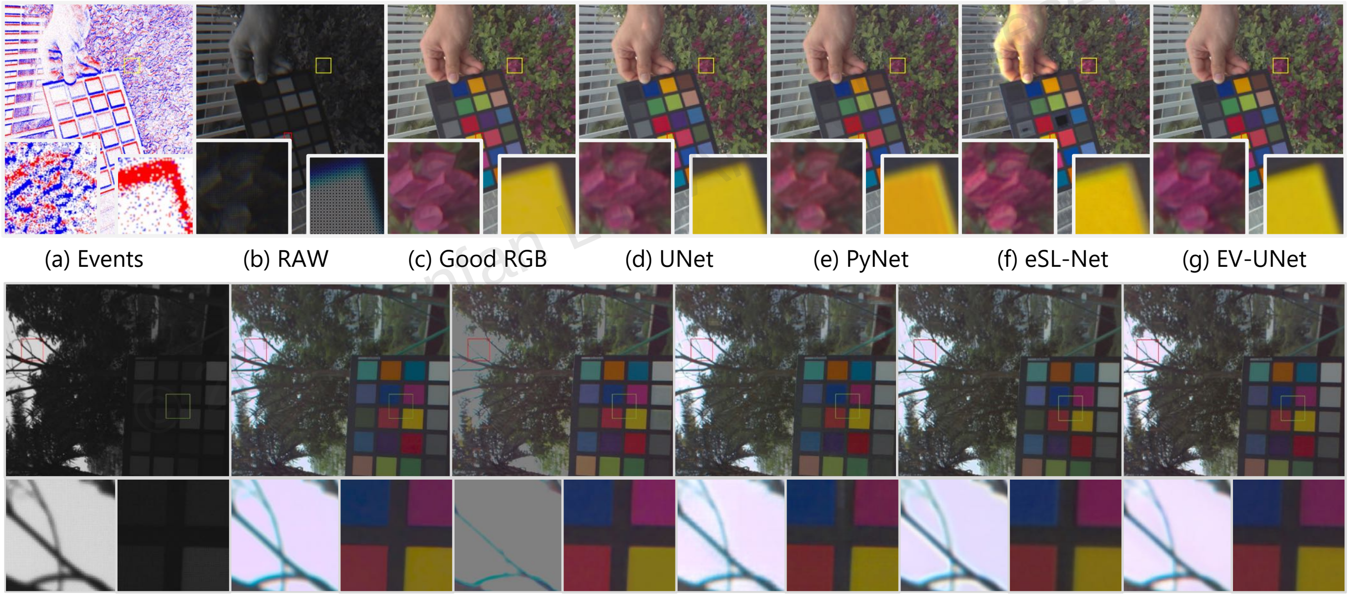


(a) CIEDE00 Error (b) CIELAB Δ_{ab} Error (c) Time Domain Fluctuations

RAW and Color – *Benchmark*

Key Results1: Events provide clear benefits in outdoor scenes.

- Event fusion (EV-UNet) boosts UNet by +1.94 PSNR
- eSL-Net fails due to limited receptive field.



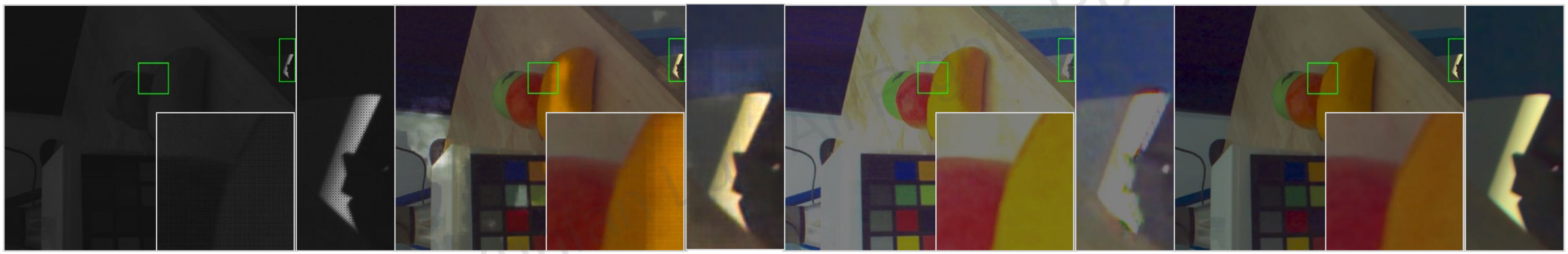
(a) Events (b) RAW (c) Good RGB (d) UNet (e) PyNet (f) eSL-Net (g) EV-UNet

(h) RAW (i) Good RGB (j) AWNet (k) Swin-Transformer (l) Unet (m) PyNetCA

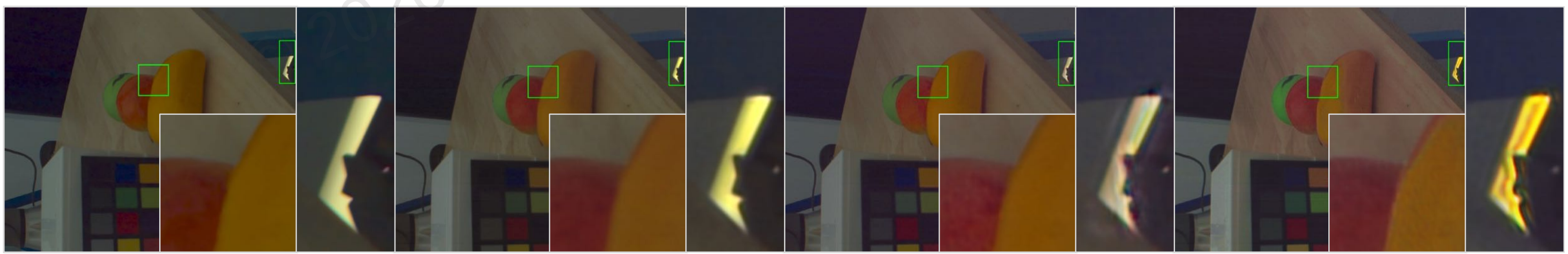
RAW and Color – *Benchmark*

Key Result 2: In indoor scenes, events may even hurt reconstruction quality.

- The benefit of events is not universal: in some indoor scenes, event guidance degrades color reconstruction.
- Flickering lights disrupt event patterns



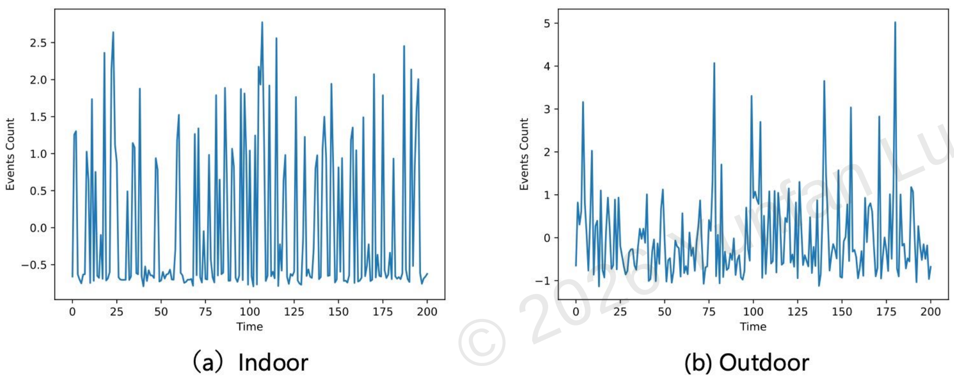
(a) Raw (b) Swin-Transformer (c) AWNet (d) MW-ISPNet



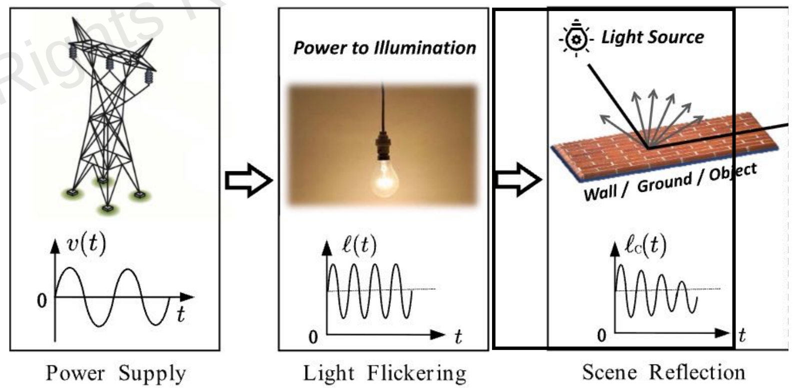
(e) Good RGB (f) U-Net (g) PyNetCA (h) PyNet

RAW and Color – *Benchmark*

In-depth analysis: Flickering illumination can generate spurious events that do not reflect scene dynamics, but instead encode sensor-light interactions, which must be explicitly considered in event-guided ISP.



(a) Indoor scenes exhibit much denser event responses than outdoor scenes, suggesting additional event triggers beyond true scene motion.



(b) Indoor light flicker can propagate through scene reflection and induce spurious events, introducing sensor-light interactions that are irrelevant to the desired ISP signal.

RAW and Color – *Summary*

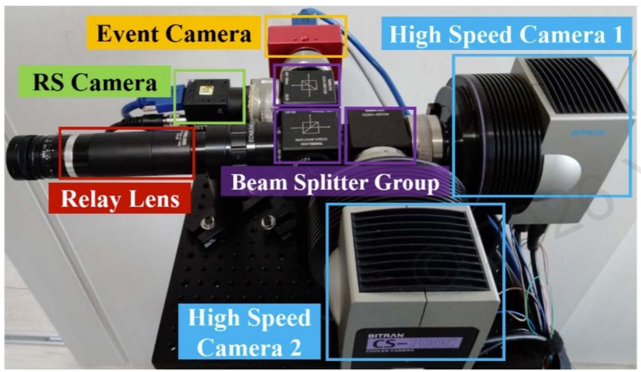
1. Clearer question formulation + physics-aware modeling = simpler and more efficient image reconstruction.
2. Better motion representation + more tailored architecture design = stronger continuous space-time reconstruction.
3. New dataset construction + new task formulation = a broader event-based imaging paradigm.
4. RAW-domain exploration + in-depth analysis = a new direction for event-guided ISP.

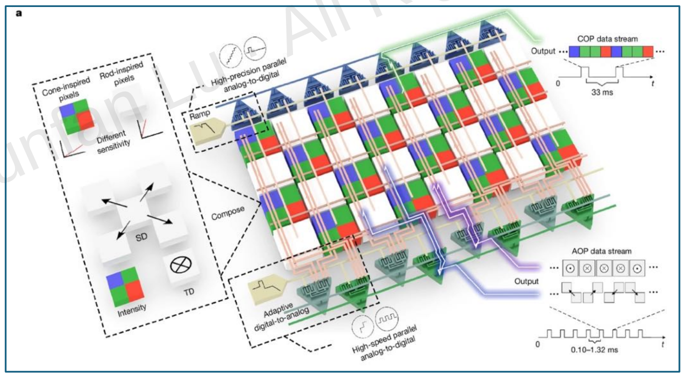
Five Representative Works

1. **Temporal Reconstruction**: UniINR: Event-Guided Unified Rolling Shutter Correction, Deblurring, and Interpolation. ECCV 2024
2. **Spatial-Temporal Reconstruction**: Continuous Space-Time Video Super-Resolution via Event Camera, IJCV 2025
3. **Dynamic Range Recovery**: SEE: See Everything Every Time, IJCV 2025
4. **RAW and Color Imaging**: RGB-Event ISP: Dataset, Benchmark, and Direction, ICLR 2025
5. **Sensor Modeling and Simulation**: Hybrid Event-Frame Sensors: Modeling, Calibration, and Simulation

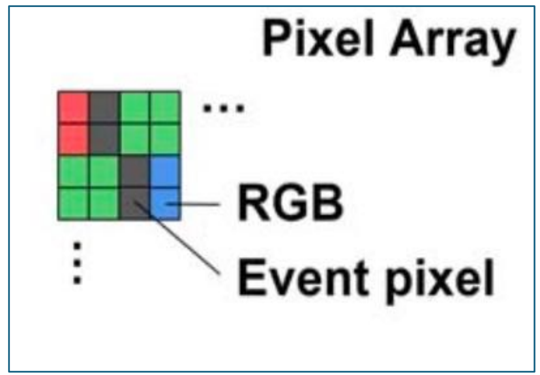
Sensor Modeling - *Background*

- Key observation: Hybrid event-frame sensors integrate APS and EVS on one chip, combining dense intensity with low-latency events.
- This tighter integration also introduces more complex noise patterns than conventional cameras.
- **Can we use statistical models to describe the event sensor?**





Nature 2024



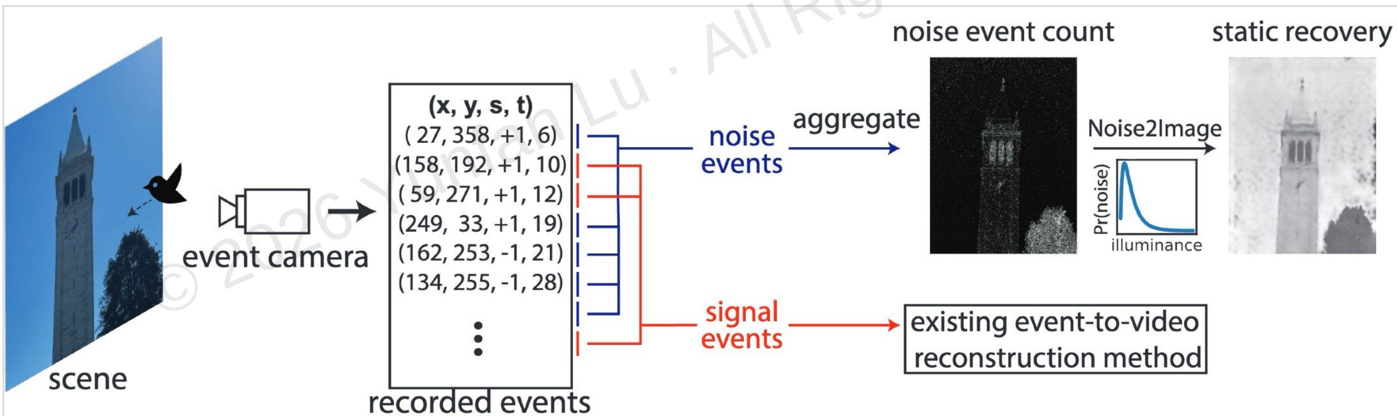
ISSCC 2023 (SONY)

Optical-level Hybrid

Sensor-level Hybrid

Sensor Modeling – *Related Works*

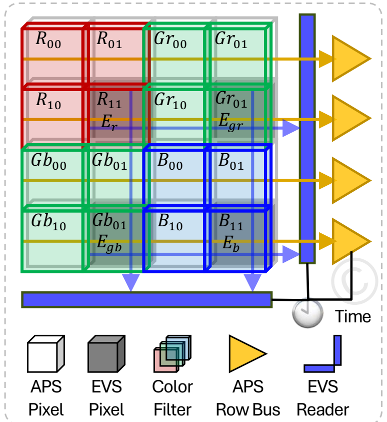
- Standard camera noise: APS noise modeling is well studied but focuses on frame sensors alone [1].
- Event camera noise: Event noise was observed to be related to illumination [2], but joint noise modeling was not modeled.
- **APS and EVS should be modeled from the shared physical signal.**



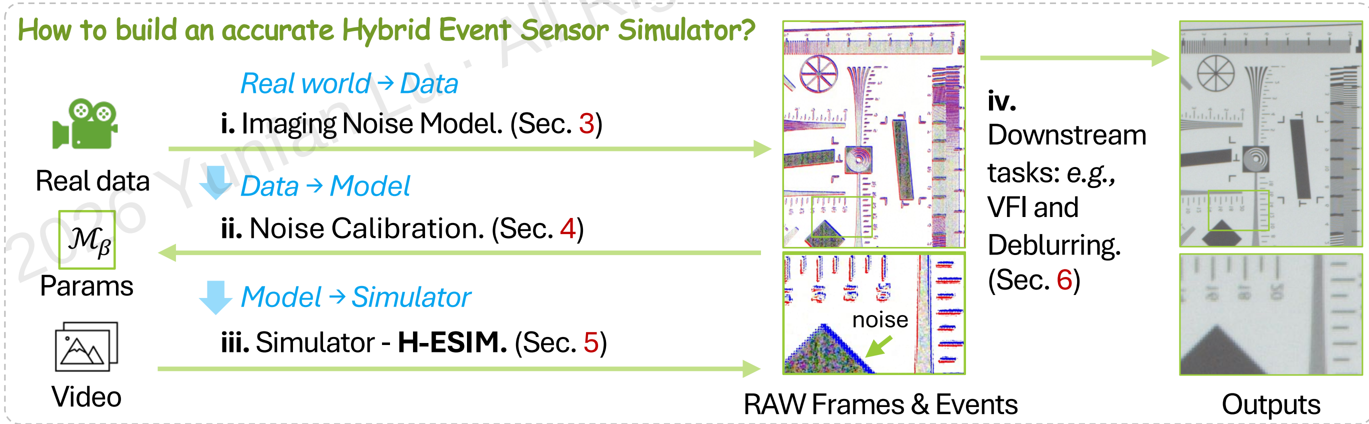
[1] Wei, Kaixuan, Ying Fu, Jiaolong Yang, and Hua Huang. "A physics-based noise formation model for extreme low-light raw denoising." CVPR 2020.
[2] Cao, Ruiming, Dekel Galor, Amit Kohli, Jacob L. Yates, and Laura Waller. "Noise2Image: noise-enabled static scene recovery for event cameras." Optica 12, no. 1 (2025): 46-55.

Sensor Modeling – *Key Insight*

- Key insight: APS and EVS share the same optical input and should be modeled within one physics-informed statistical framework.
  - APS: performs integral sampling and outputs RAW intensity frames.
  - EVS: performs differential thresholder readout and outputs events.
- **The next step is to build a full pipeline: model → calibration → simulator.**



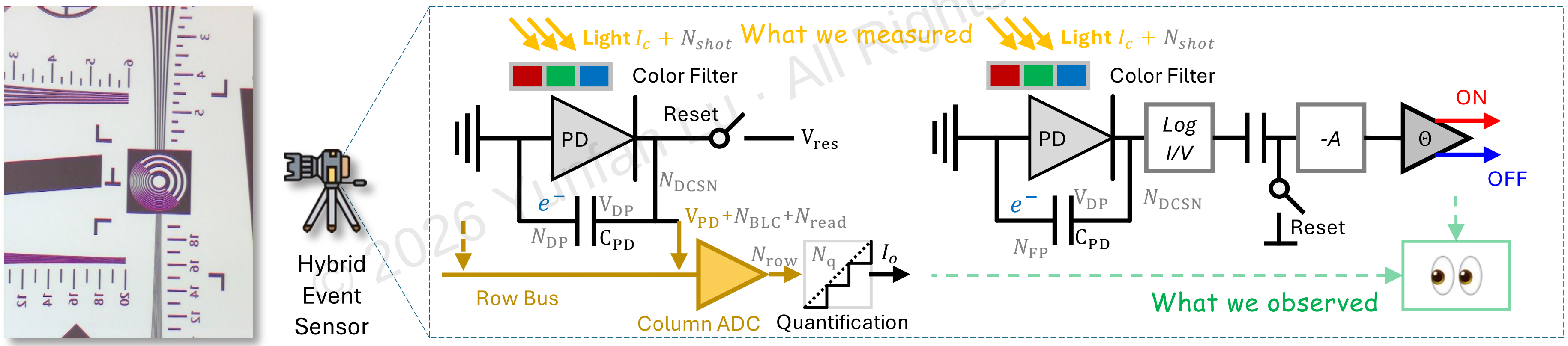
(a) Hybrid Event Sensor



(b) Overview of Our Framework

Sensor Modeling - *Modeling*

- Noise describes the gap between the **observed signal** and the **true signal**.
- Noise exists in all modalities and is unavoidable.
- Hybrid event sensors share the same optical system with input I_c for APS and EVS.
- **Complex noise can be represented as a combination of a series of simple noises.**

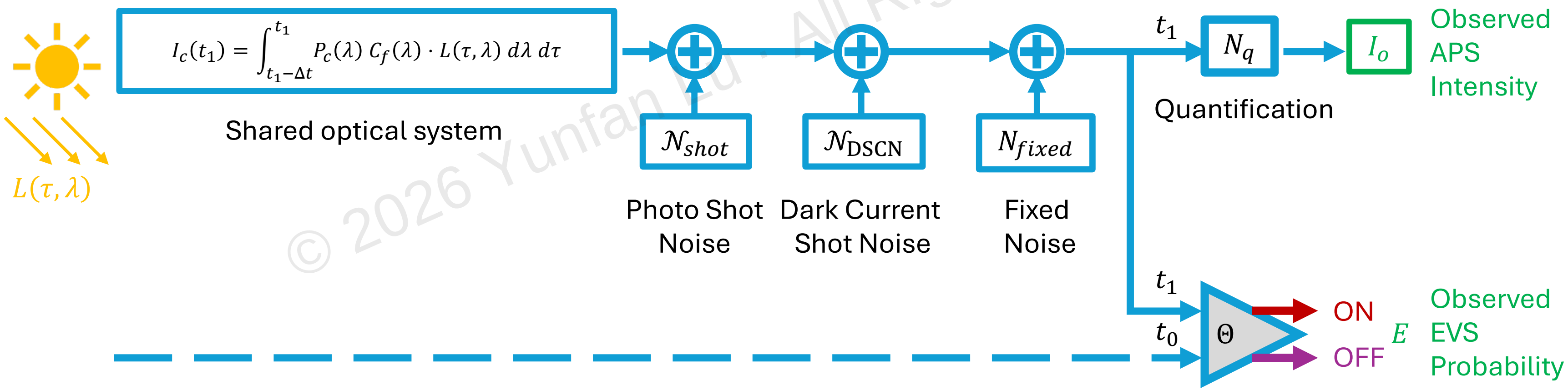


(a) Real World (b) Camera (c) APS Pixel Simplified Circuit Diagram (d) EVS Pixel Simplified Circuit Diagram

Sensor Modeling - *Modeling*

Noise can be grouped into three categories, which are shared by both APS and EVS:

- Photon-related noise (shot noise, illumination dependent)
- Dark current noise (exposure and temperature dependent)
- Fixed noise from circuit process (e.g., black-level correction noise, readout, FPN, quantization)



Sensor Modeling - *Modeling*

Activate Pixel Sensor

- $I_o = I_c + N_{shot}(I_c) + N_{time}(\Delta t) + N_{fixed}$

Photo Shot Noise:	• $N_{shot}(I_c) = \mathcal{N}(0, \sigma^2_{shot}) = \mathcal{N}(0, \beta_1 \cdot I_c)$
Hot pixel and dark-current shot noise	• $N_{time}(\Delta t) = \Delta t \cdot N_{DP} + \mathcal{N}(0, \sigma^2_{DCSN}) = \mathcal{N}(0, \beta_2 \cdot \Delta t)$
Fixed noise:	• $N_{fixed} = N_{FP} + \mathcal{N}(0, \sigma^2_{read}) + \mathcal{U}(\frac{-1}{2q}, \frac{1}{2q})$

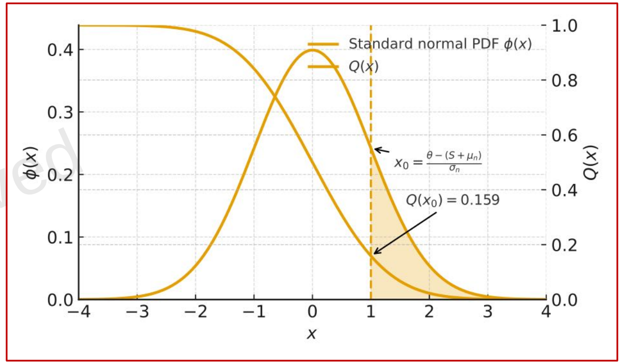
- APS noise N_a:
  - $N_a = I_o - I_c = \mathcal{N}(0, \sigma^2_{shot} + \sigma^2_{DCSN} + \sigma^2_{read} - \sigma^2_{\times}) + \Delta t \cdot N_{DP} + N_{FP} + N_q$
- Its variance can be expressed as a polynomial of intensity and exposure time.
- $\mathrm{Var}(N_a) = \sigma^2_{shot} + \sigma^2_{DCSN} + \sigma^2_{read} - \sigma^2_{\times} \approx \beta_0 + \beta_1 \cdot I_c + \beta_2 \cdot \Delta t + \beta_3 \cdot I_c^2 + \beta_4 \cdot I_c \cdot \Delta t + \beta_5 \cdot \Delta t^2$
- **APS noise can be described as a superposition of a series of Gaussian and uniform distributions.**

Sensor Modeling - *Modeling*

Event Vision Sensor Noise

Noise part, N_e **Signal part S**

- $V_{t_1} = \hat{V}_{t_1} + \boxed{N_{shot} + N_{DCSN} + N_{FP} + V_{PD}}$
- $\Delta V_{t_1} = \log(V_{t_1}) - \log(V_{t_0}) = \boxed{log\left(1 + \frac{n_{t_1}}{\widehat{V_{t_1}} + V_{PD}}\right) - log\left(1 + \frac{n_{t_0}}{\widehat{V_{t_0}} + V_{PD}}\right)} + \boxed{\log\left(\frac{\widehat{V_{t_1}} + V_{PD}}{\widehat{V_{t_0}} + V_{PD}}\right)}$
  - Taylor expansion, $n_{t_1} \ll \widehat{V_{t_1}} + V_{PD}$
- Noise part: $N_e = log\left(1 + \frac{n_{t_1}}{\widehat{V_{t_1}} + V_{PD}}\right) - log\left(1 + \frac{n_{t_0}}{\widehat{V_{t_0}} + V_{PD}}\right) \approx \frac{n_{t_1}}{\widehat{V_{t_1}} + V_{PD}} - \frac{n_{t_0}}{\widehat{V_{t_0}} + V_{PD}}$
  - $\boldsymbol{V}_{t_1} = \widehat{V_{t_1}} + V_{PD}, \boldsymbol{V}_{t_0} = \widehat{V_{t_0}} + V_{PD}$
- $N_e \sim \mathcal{N}\left(N_{FP}\left(\frac{1}{\boldsymbol{V}_{t_1}} - \frac{1}{\boldsymbol{V}_{t_0}}\right), 2\sigma^2_{\text{shot}} + 2\sigma^2_{\text{DCSN}} - 2\rho\sigma_{\text{shot}}\sigma_{\text{DCSN}}\left(\frac{1}{\boldsymbol{V}_{t_1}{}^2} + \frac{1}{\boldsymbol{V}_{t_0}{}^2}\right)\right)$
  - $\mu_n = N_{FP}\left(\frac{1}{\boldsymbol{V}_{t_1}} - \frac{1}{\boldsymbol{V}_{t_0}}\right), \sigma_n^2 = 2\sigma^2_{\text{shot}} + 2\sigma^2_{\text{DCSN}} - 2\rho\sigma_{\text{shot}}\sigma_{\text{DCSN}}\left(\frac{1}{\boldsymbol{V}_{t_1}{}^2} + \frac{1}{\boldsymbol{V}_{t_0}{}^2}\right)$
- **Key Insight**: We use $Q -$ function to establish the relationship between N_e and observed events.



$Q -$ function

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{t^2}{2}} dt$$

Sensor Modeling - *Modeling*

Noise Event Observation Probability, $\widehat{V_{t_1}} + V_{PD}$ is a static value.

- $P_+ = \mathbf{P}(\Delta V_k > \theta) = Q\left(\frac{\theta - \mu_n}{\sigma_n}\right)$
- $P_- = \mathbf{P}(\Delta V_k < -\theta) = 1 - Q\left(\frac{-\theta + \mu_n}{\sigma_n}\right)$
- $P = \frac{P_+ + P_-}{2}$
- $\sigma_n^2 = \frac{(2\sigma_{\text{shot}}^2 + 2\sigma_{\text{DCSN}}^2 - 2\rho\sigma_{\text{shot}}\sigma_{\text{DCSN}})}{(\widehat{V_{t_1}} + V_{PD})^2} = \left(\frac{\theta}{Q^{-1}(P)}\right)^2$
- $\boxed{Q^{-1}(P)} = \frac{\theta\left(\boxed{\widehat{V_{t_1}}} + V_{PD}\right)}{\sqrt{(2\sigma_{\text{shot}}^2 + 2\sigma_{\text{DCSN}}^2 - 2\rho\sigma_{\text{shot}}\sigma_{\text{DCSN}})}}$

We observed   Controllable variables

1. $\theta = \beta_0 \cdot \Theta$, Θ is the hardware threshold (0.75 mV).
2. $\widehat{V_{t_1}} + V_{PD} = \beta_1 I_c + \beta_2$
3. $\sigma_{\text{shot}} = \beta_3 I_c$
4. $\sigma_{\text{DCSN}} = \beta_4$
5. $\rho = \beta_5$

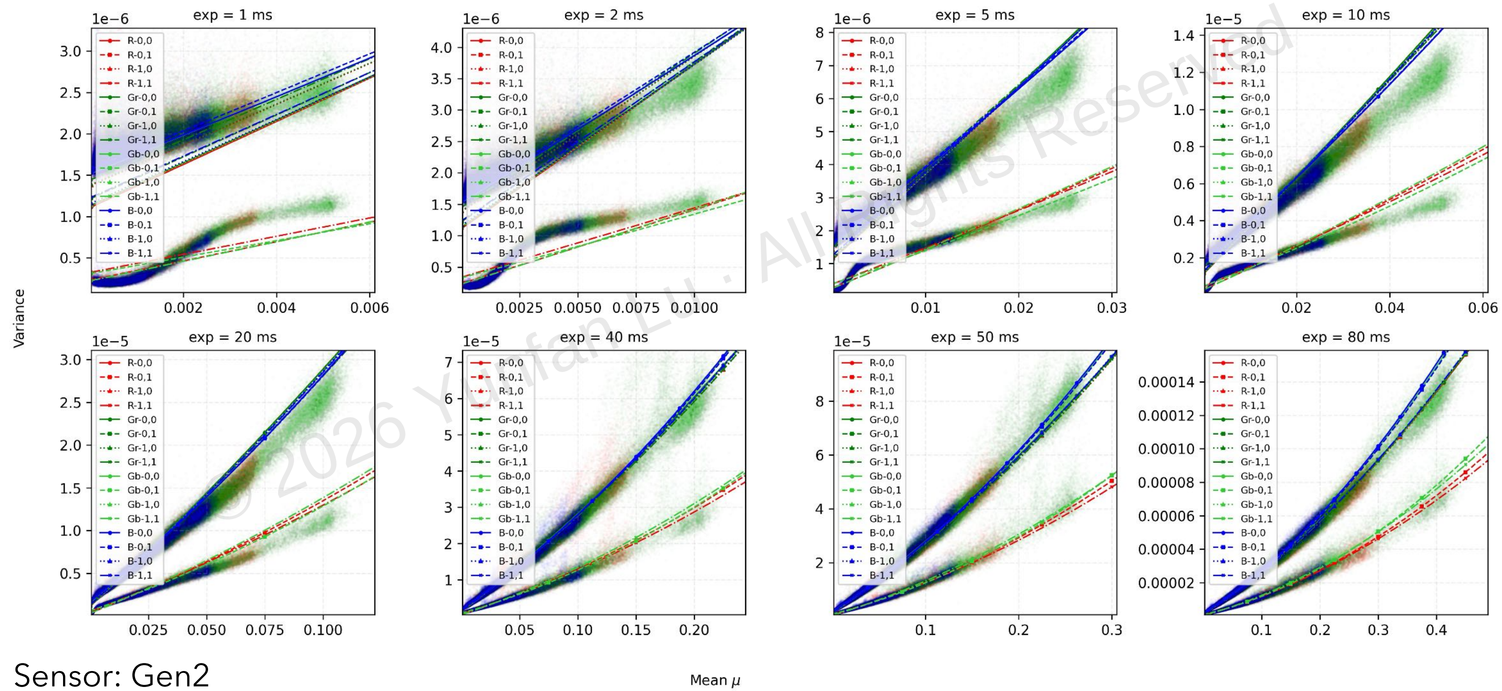
$$Q^{-1}(P) = \frac{\beta_0 \cdot \Theta\,(\beta_1 I_c + \beta_2)}{\sqrt{2(\beta_3 I_c)^2 + 2\beta_4^2 + 2\beta_5\beta_3 I_c \beta_4}}$$

The noise of EVS can be described using Q-Function.

https://yunfanlu.github.io/

Sensor Modeling - *Calibration*

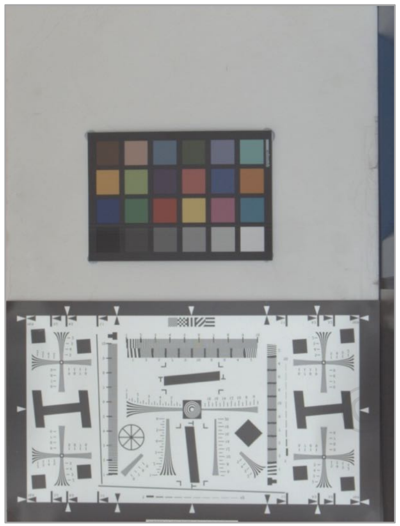
- APS calibration validates that the proposed noise model matches real RAW noise across exposure settings.

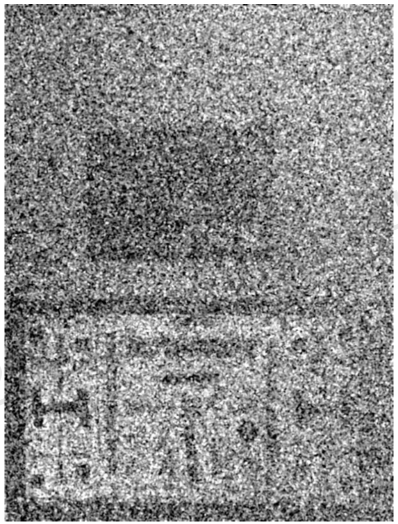


Sensor: Gen2

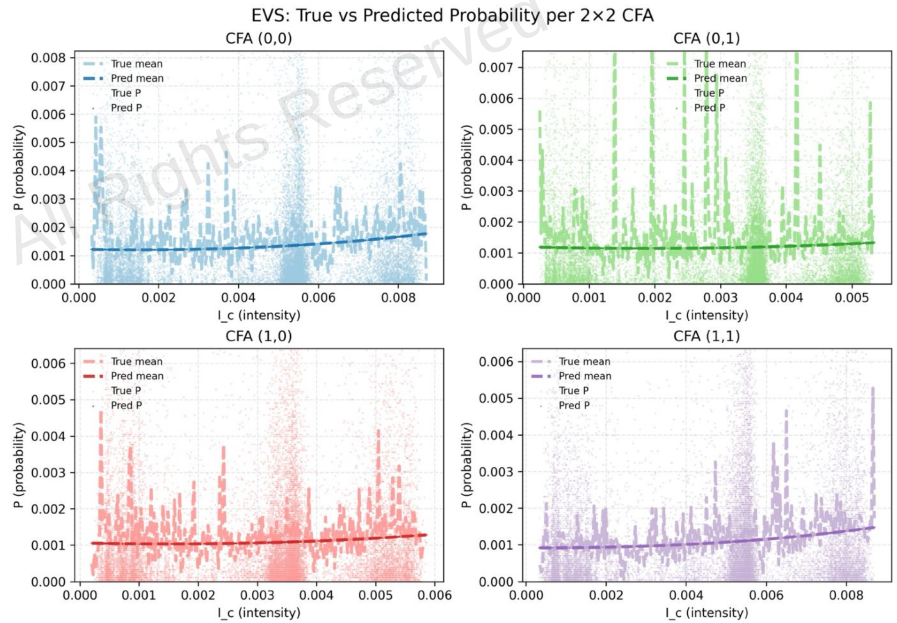
Sensor Modeling - *Calibration*

- The calibrated noise distribution of EVS explains the relationship between light intensity and the actual noise distribution.



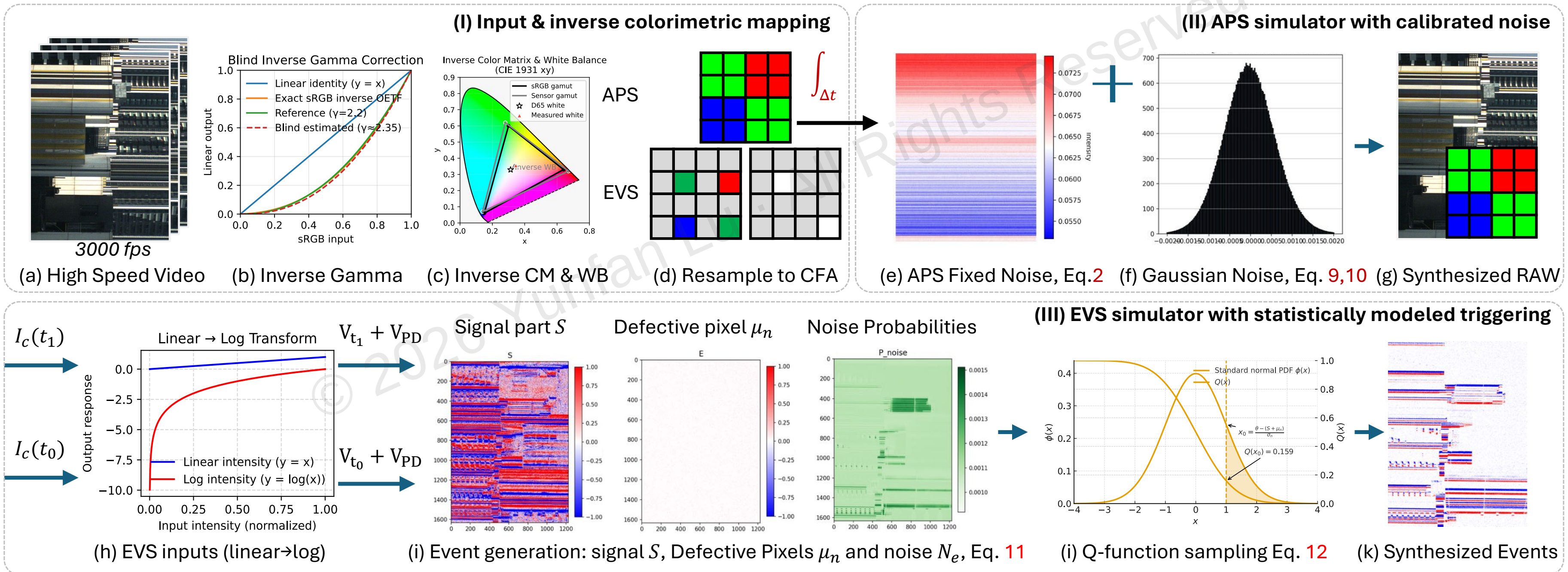


Sensor: Eiger



Sensor Modeling - *Simulation*

- With calibrated APS and EVS noise, we build H-ESIM – a hybrid sensor simulator for RAW frames and events.
- This enables realistic synthetic data for downstream tasks.



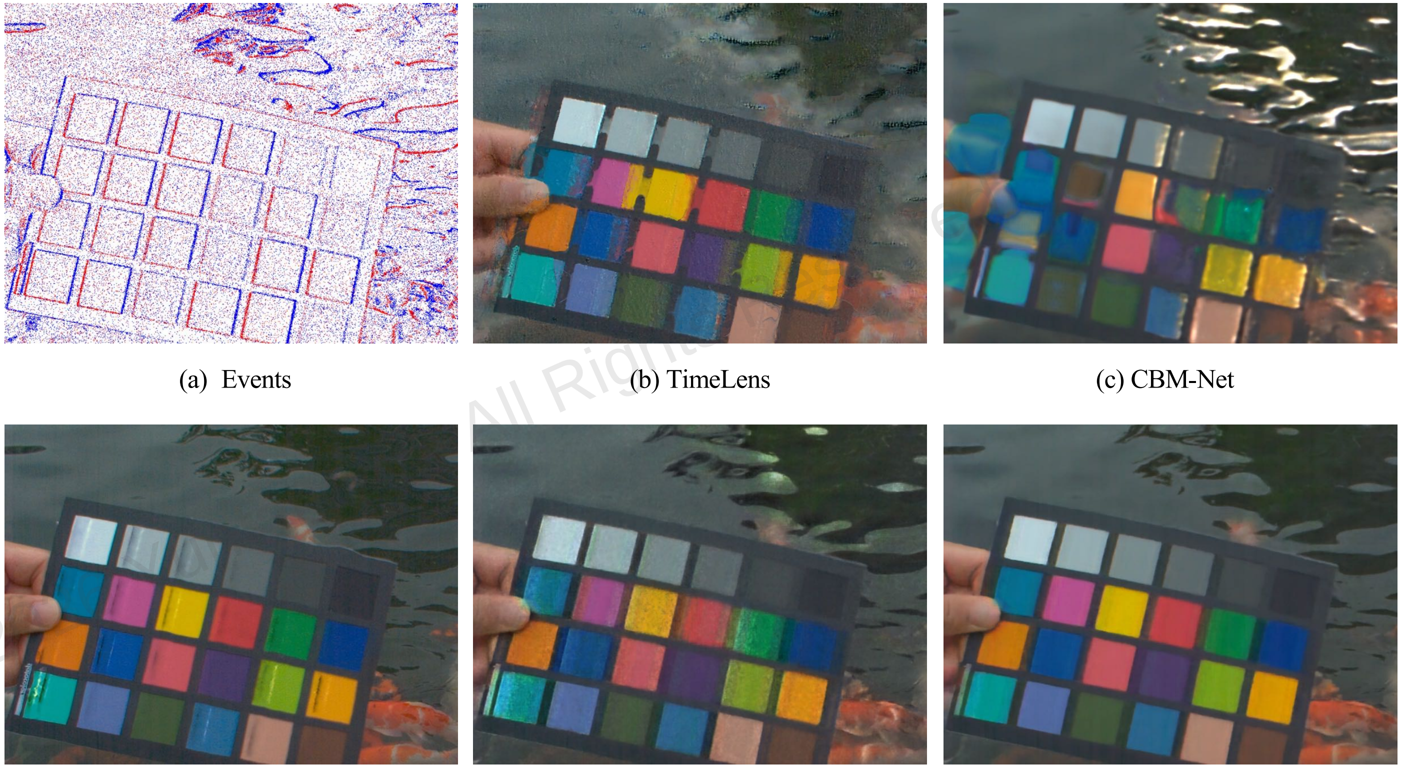
(a) High Speed Video (b) Inverse Gamma (c) Inverse CM & WB (d) Resample to CFA (e) APS Fixed Noise, Eq.2 (f) Gaussian Noise, Eq. 9,10 (g) Synthesized RAW

(h) EVS inputs (linear→log) (i) Event generation: signal S, Defective Pixels μ_n and noise N_e, Eq. 11 (i) Q-function sampling Eq. 12 (k) Synthesized Events

Sensor Modeling - *Experiments*

Results 1: H-ESIM improves interpolation quality on real hybrid-sensor data.

- After fine-tuning with H-ESIM, reconstructed frames are sharper and more coherent.



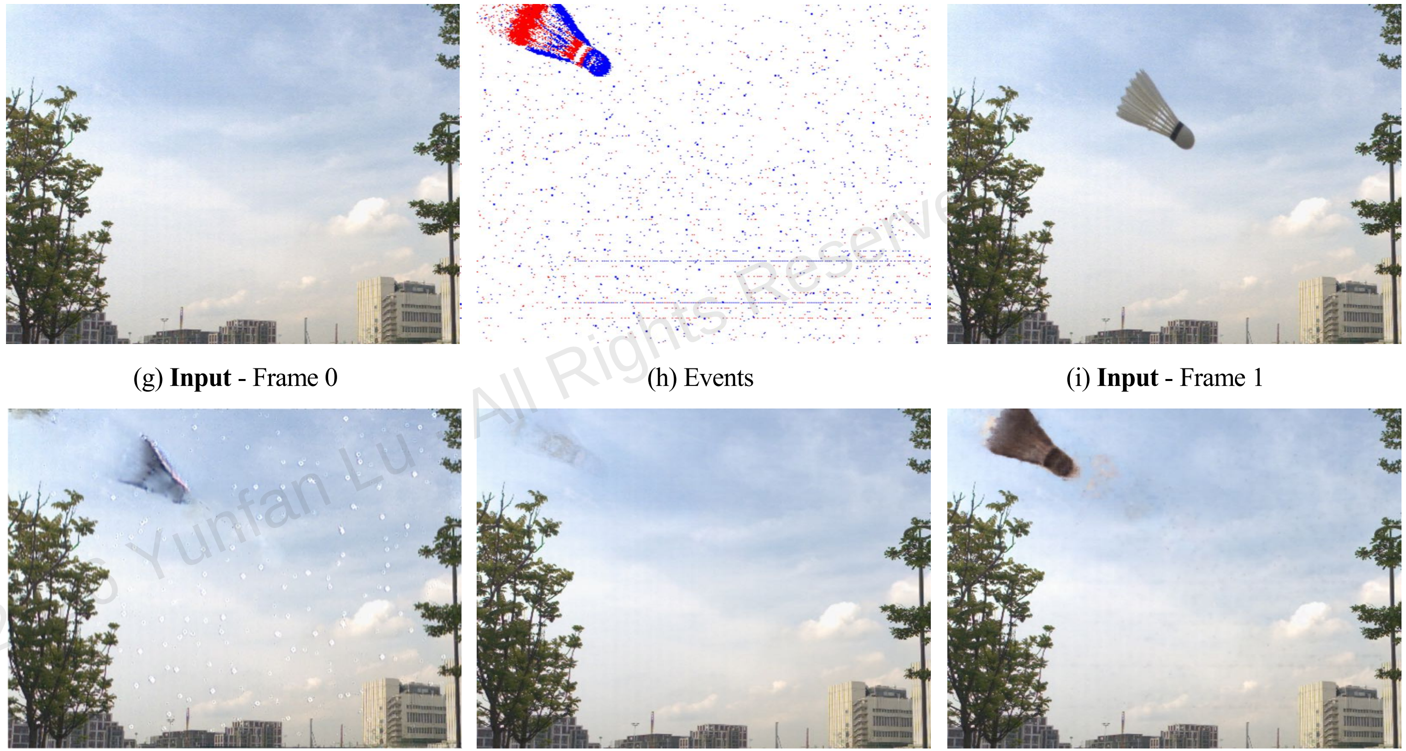
(a) Events (b) TimeLens (c) CBM-Net

(d) TimeLens-XL (e) HRINR-*wo* (f) HRINR-*w*

Sensor Modeling - *Experiments*

Results 1: H-ESIM improves interpolation quality on real hybrid-sensor data.

- After fine-tuning with H-ESIM, reconstructed frames are sharper and more coherent.



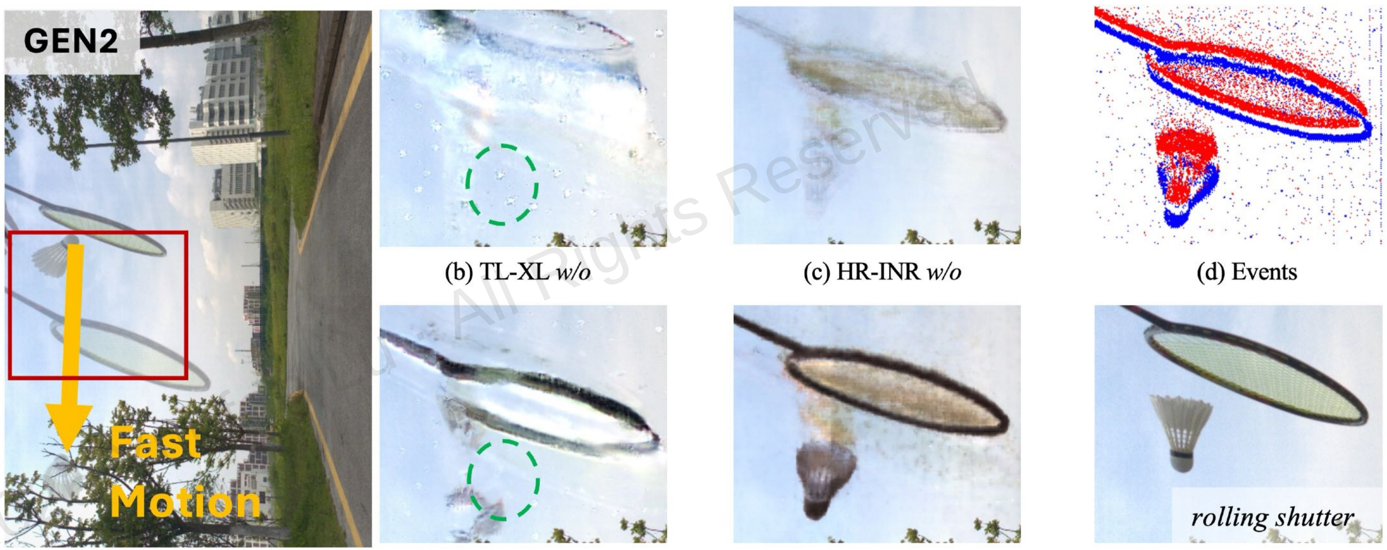
(g) **Input** - Frame 0 (h) Events (i) **Input** - Frame 1

(j) TimeLens-XL (k) HRINR-*wo* (l) HRINR-*w*

https://yunfanlu.github.io/

Sensor Modeling - *Experiments*

Reference-based interpolation confirms that H-ESIM reduces the synthetic–real gap.



(a) Inputs (b) TL-XL *w/o* (c) HR-INR *w/o* (d) Events (e) TL-XL *w* (f) HR-INR *w* (g) Reference Frame

Sensor Modeling – *Summary*

1. Clearer question formulation + physics-aware modeling = simpler and more efficient image reconstruction.
2. Better motion representation + more tailored architecture design = stronger continuous space-time reconstruction.
3. New dataset construction + new task formulation = a broader event-based imaging paradigm.
4. RAW-domain exploration + in-depth analysis = a new direction for event-guided ISP.
5. Physical modeling + calibrated parameters + realistic simulation = a new way for hybrid event-frame imaging.

Thank You!
Q&A

ylu066@connect.hkust-gz.edu.cn