



WeLight@HKU



Seeing Beyond RGB

Neural Optics for Stereo RGBD and Hyper-spectral Imaging

Yifan (Evan) Peng

HKU EEE & CS

Apr 2025



Images from public domain



**Applications are
different, imagers
& ISPs are not**



For Education Use Only, 2025
by E.Y. PENG@HKU



Co-design of Optics, Sensing, and Algorithms



Ray optics → Wave optics

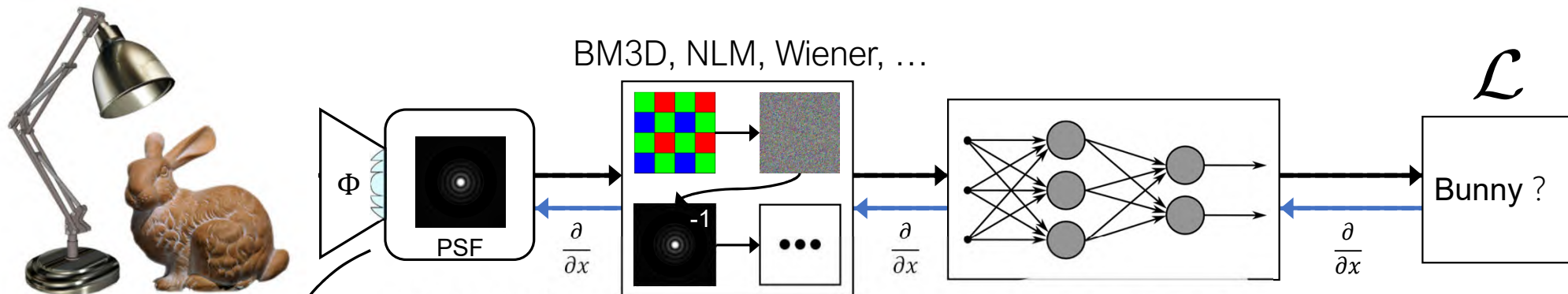
Diffractive Optical Elements

- thin & lightweight structure
- large & flexible design space
- availability of mass fabrication
- color dispersion (downside)



Promising platform for imaging & visual computing apps.

Neural Optics for Domain-specific Apps. Imaging

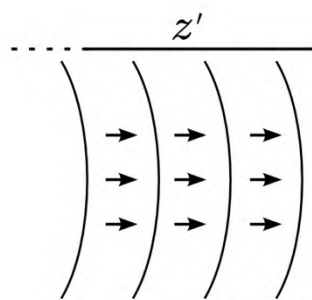


Wavefront coding,
Diffractive Achromat

Differentiable Lenses

Differentiable Image Processing

Optimize end-to-end via an intermediate term



$$\exp(jk\sqrt{x^2 + z'^2})$$

Wave optics PSF simulator

- Performance & robustness gains
- Domain-specific hardware w/ reduced cost
- New design space: The “*BunnyCam*”

* End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging, V Sitzmann*, S Diamond*, Y Peng*, et al. – ACM TOG, 2018.

Differentiable Cameras for Domain-specific Apps.

End-to-end Optimization of Optics and Image Processing for Achromatic Extended Depth of Field and Super-resolution Imaging

VINCENT SITZMANN*, Stanford University
 STEVEN DIAMOND*, Stanford University
 YIFAN PENG*, The University of British Columbia and Stanford University
 XIONG DUN*, King Abdullah University of Science and Technology
 STEPHEN BOYD, Stanford University
 WOLFGANG HEIDRICH, King Abdullah University of Science and Technology
 FELIX HEIDE, Stanford University
 GORDON WETZSTEIN, Stanford University



Single-shot Extended DOF
 ACM TOG '18

Deep Optics for Single-shot High-dynamic-range Imaging

Christopher A. Metzler Hayato Ikoma Yifan Peng Gordon Wetzstein
 [metzler, ikoma, ewspeng, gordon.wetzstein}@stanford.edu

Abstract

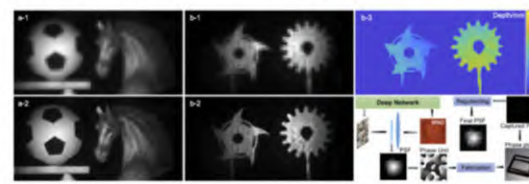
High-dynamic-range (HDR) imaging is crucial for many applications. Yet, acquiring HDR images with a single shot remains a challenging problem. Whereas modern deep learning approaches are successful at hallucinating plausible HDR content from a single low-dynamic-range (LDR) image, saturated scene details often cannot be faithfully recovered. Inspired by recent deep optical imaging approaches, we interpret this problem as jointly training an optical encoder and electronic decoder where the encoder is parameterized by the point spread function (PSF) of the lens, the bottleneck is the sensor with a limited dynamic range, and the decoder is a convolutional neural network (CNN). The lens surface is then made, fabricated with the CNN as a teacher.



Single-shot HDR Imaging
 IEEE CVPR '20

End-to-end Learned, Optically Coded Super-resolution SPAD Camera

QILIN SUN, King Abdullah University of Science and Technology
 JIAN ZHANG, Peking University
 XIONG DUN, Tongji University
 BERNARD GHANEM, King Abdullah University of Science and Technology
 YIFAN PENG, Stanford University
 WOLFGANG HEIDRICH, King Abdullah University of Science and Technology



Super-resolution SPAD Camera
 ACM TOG '20; IEEE CVPR '18

Learned Large Field-of-View Imaging With Thin-Plate Optics

YIFAN PENG*, Stanford University
 QILIN SUN*, King Abdullah University of Science and Technology
 XIONG DUN*, King Abdullah University of Science and Technology
 GORDON WETZSTEIN, Stanford University
 WOLFGANG HEIDRICH, King Abdullah University of Science and Technology
 FELIX HEIDE, Princeton University



Large FOV & Lightweight
 OSA OE '15; ACM TOG '19

Depth from Defocus with Learned Optics for Imaging and Occlusion-aware Depth Estimation

Hayato Ikoma, Cindy M. Nguyen, Christopher A. Metzler, Member, IEEE, Yifan Peng, Member, IEEE, and Gordon Wetzstein, Senior Member, IEEE

Abstract—Monocular depth estimation remains a challenging problem, despite significant advances in neural network architectures that leverage potential depth cues alone. Inspired by depth from defocus and emerging point spread function engineering approaches that optimize programmable optics and to end with depth estimation networks, we propose a new and improved framework for depth estimation from a single RGB image using a learned phase-coded aperture. Our optimized aperture code uses rotational symmetry constraints for computational efficiency, and we pair them with the optics and the network using an occlusion-aware image formation model that provides more accurate defocus blur at depth discontinuities than previous techniques do. Using this framework and a custom prototype camera, we demonstrate state-of-the-art image and depth estimation quality among end-to-end optimized computational cameras in simulation and experiment.

Index Terms—Computational Photography, Computational Optics

AiF & Depth Estimation
 IEEE ICCP '21

Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging

XIONG DUN,^{1,2} HAYATO IKOMA,¹ GORDON WETZSTEIN,¹ ZHANGSHAN WANG,^{1,2} XINBIN CHENG,^{1,2,3} AND YIFAN PENG^{1,2}

¹Institute of Precision Optical Engineering, School of Physics Science and Engineering, Tongji University, Shanghai 200092, China
²Electrical Engineering Department, Stanford University, Stanford, California 94305, USA
³MOE Key Laboratory of Advanced Micro-Structured Materials, Shanghai 200092, China
 *e-mail: chengxb@tongji.edu.cn
 *e-mail: ewspeng@stanford.edu



Full-spectrum & Lightweight
 OSA Optica '20; ACM TOG '16

Differentiable Compound Optics and Processing Pipeline Optimization for End-to-end Camera Design

ETHAN TSENG*, Princeton University, United States
 ALI MOSLEH* and FAHIM MANNAN*, Algolux, Canada
 KARL ST. ARNAUD, AVINASH SHARMA, and YIFAN PENG, Algolux, Canada
 ALEXANDER BRAUN, Hochschule DUISBURG, Germany
 DEREK NOWROUZSAHRAI, McGill University, Canada
 JEAN-FRANCOIS LALONDE, Université Laval, Canada
 FELIX HEIDE, Princeton University, United States and Algolux, Canada



Differentiable Compound Optics
 ACM TOG '21



Deep learning multi-shot 3D localization microscopy using hybrid optical-electronic computing

HAYATO IKOMA,¹ TAKAMASA KUDO,² YIFAN PENG,¹ MICHAEL BROXTON,¹ AND GORDON WETZSTEIN^{1,*}

¹Department of Electrical Engineering, Stanford University, 350 Jane Stanford Way, Stanford, California 94305, USA
²Department of Chemical and Systems Biology, Stanford University, 209 Campus Drive, Stanford, California 94305, USA
 *Corresponding author: gordon.wetzstein@stanford.edu

Received 30 August 2021; revised 2 November 2021; accepted 3 November 2021; posted 4 November 2021; published 13 December 2021

3D Microscopy Imaging
 OSA OL '21



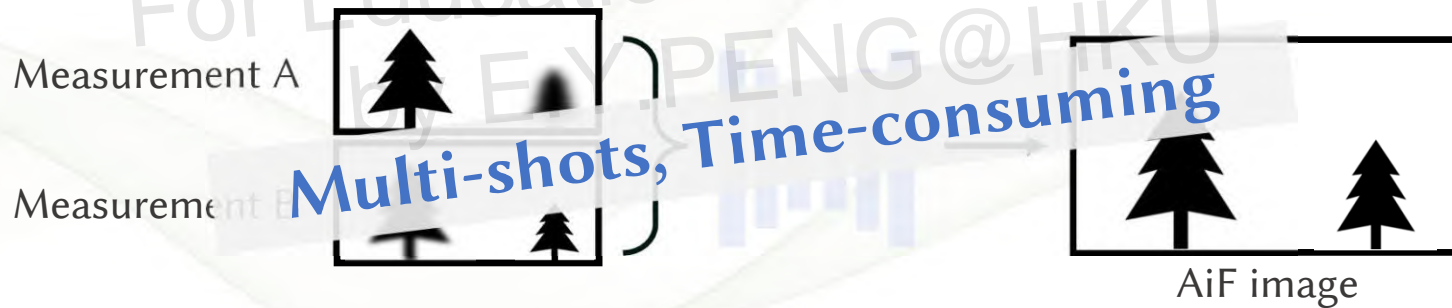
Issues in RGBD Imaging...

A. Monocular Image Reconstruction

Snapshot Depth from Defocus

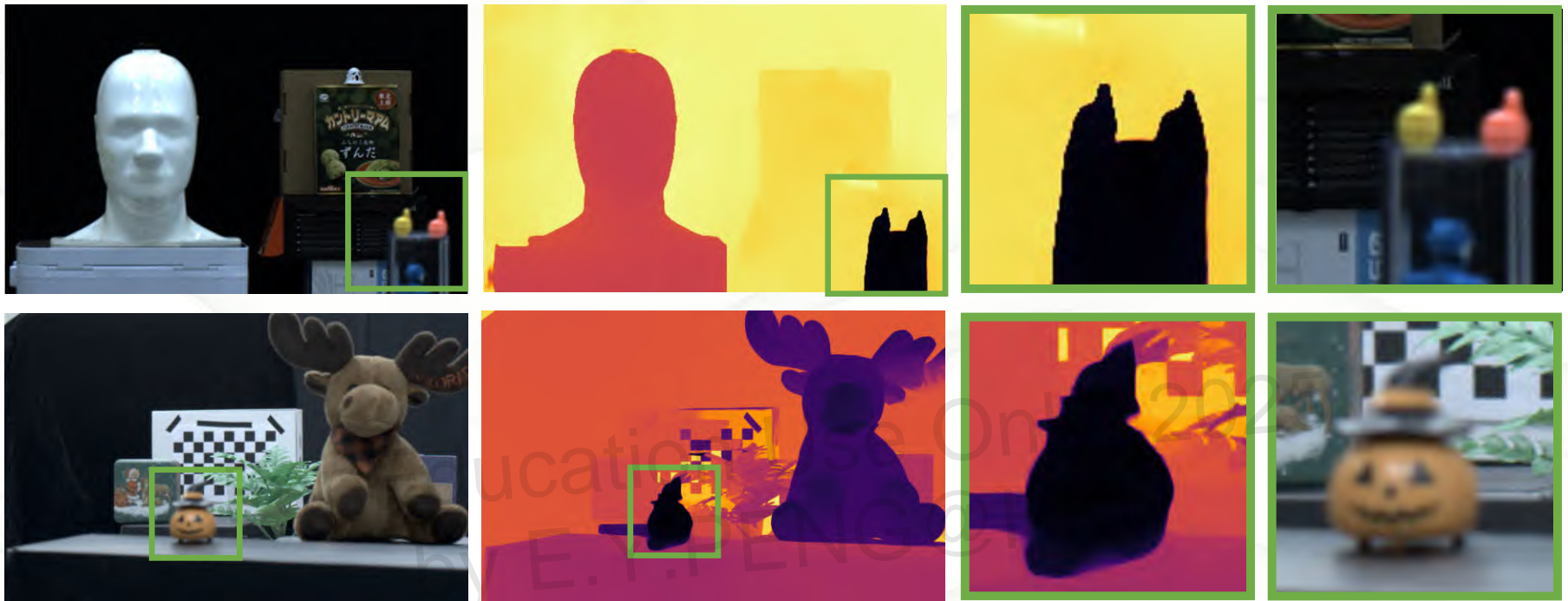


Multi-focus Image Fusion



Issues in RGBD Imaging...

B. Conventional Stereo Matching



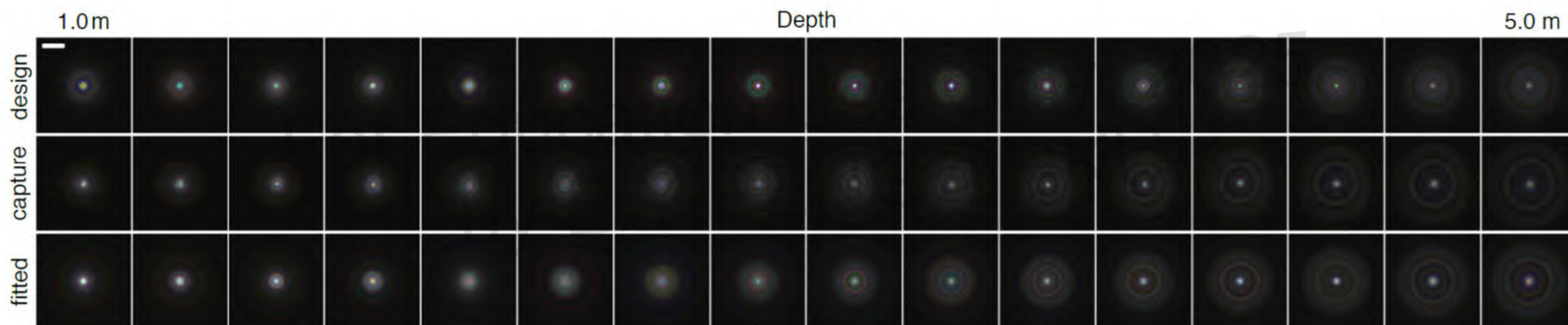
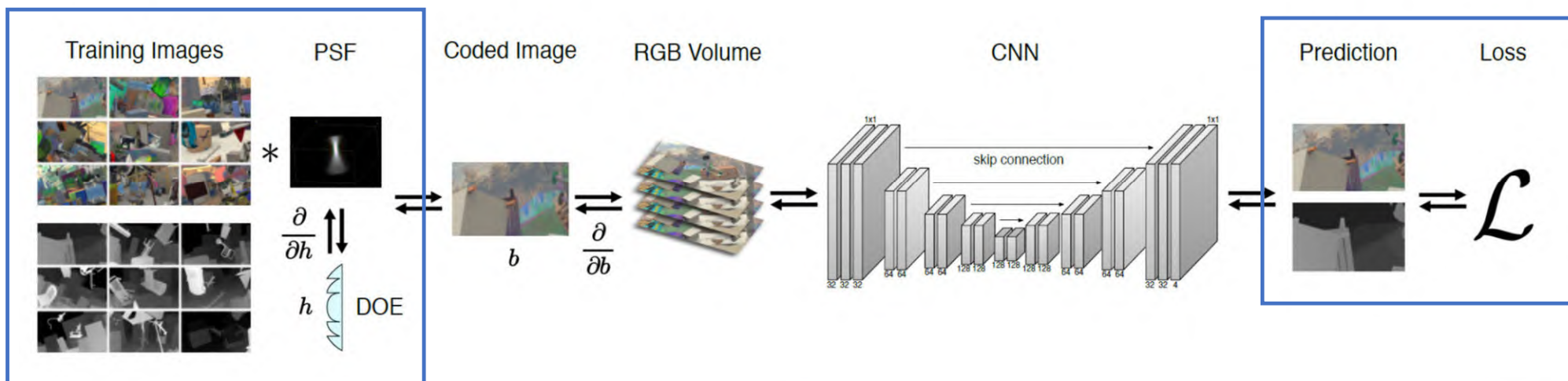
Left Measurement

Estimated disparity

Distortion caused by blur

Binocular-Encoding Optics?

Monocular Depth Estimation and All-in-focus Imaging



* Depth from Defocus with Learned Optics for Imaging and Occlusion-aware Depth Estimation, *H Ikoma, C Nguyen, C Metzler, Y Peng, G Wetzstein*– *IEEE ICCP, 2021*.

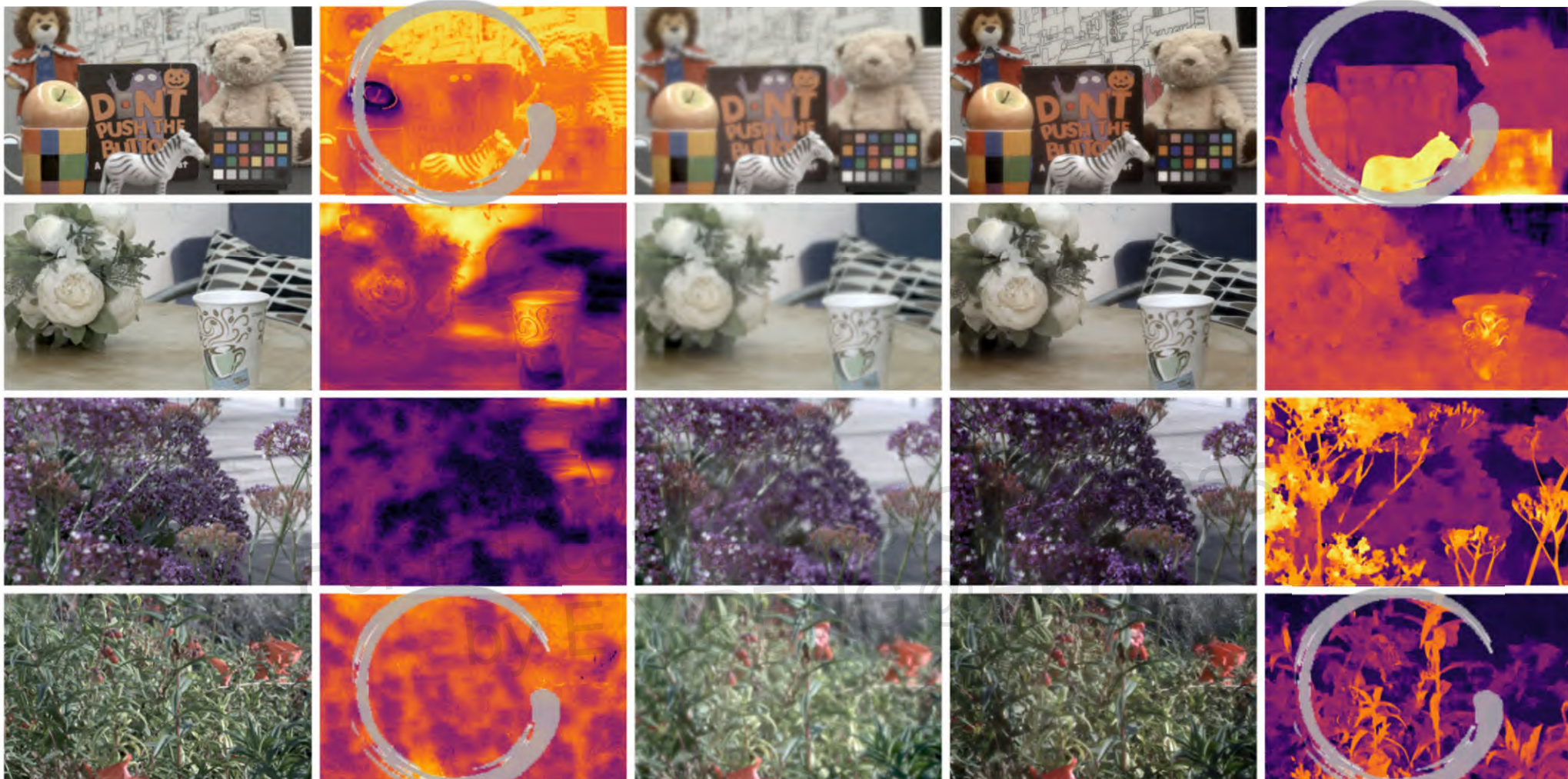
Conventional

Depth Est. (Conven.)

Ours

AiF. (Ours)

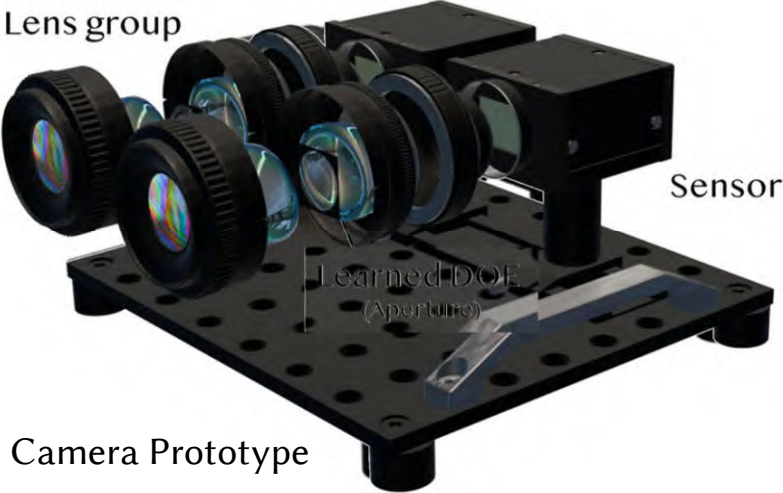
Depth Est. (Ours)



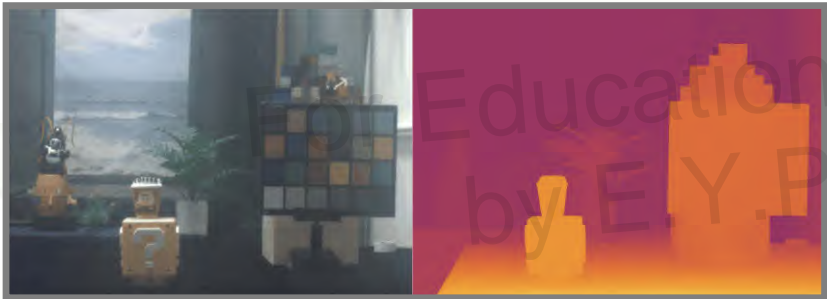
H Ikoma et al., Depth from Defocus with Learned Optics..., ICCP 2021

1.0m +5.0m

Learned Binocular-Encoding Optics for RGBD Imaging

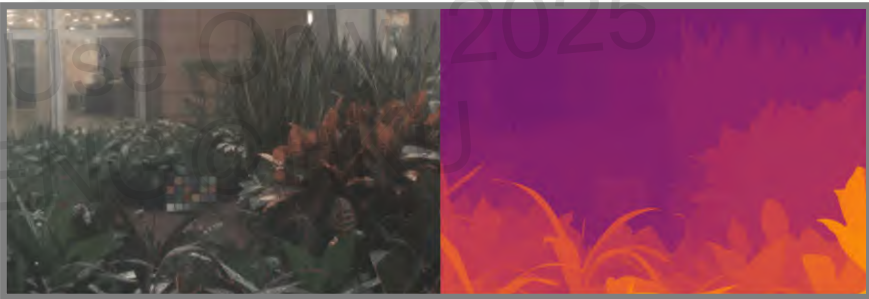


Learned Diffractive Optical Element (DOE)



Reconstructed RGB

Estimated Depth



Reconstructed RGB

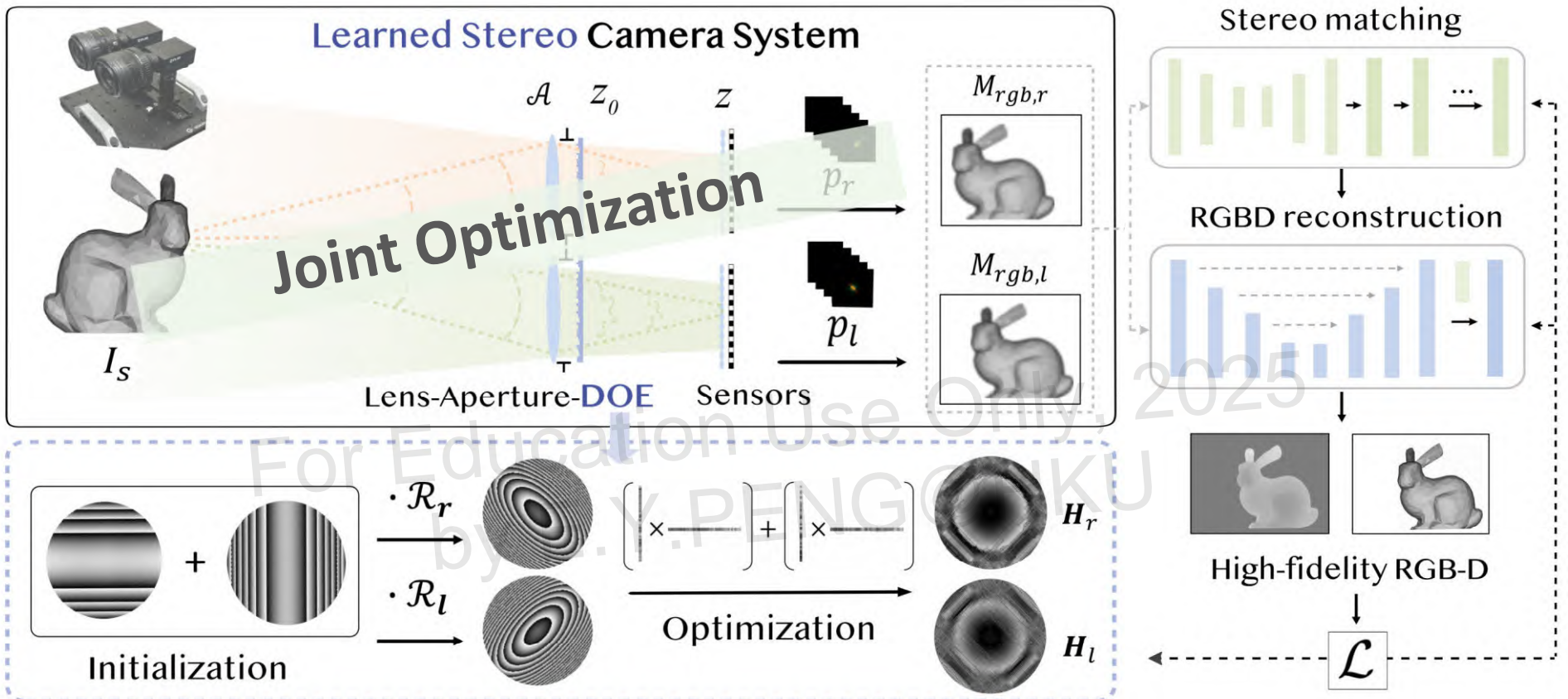
Estimated Depth

* Yuhui Liu, Liangxun Ou, Qiang Fu, QHadi Amata, Wolfgang Heidrich, Yifan Peng, *IEEE CVPR, 2025 (Oral)*

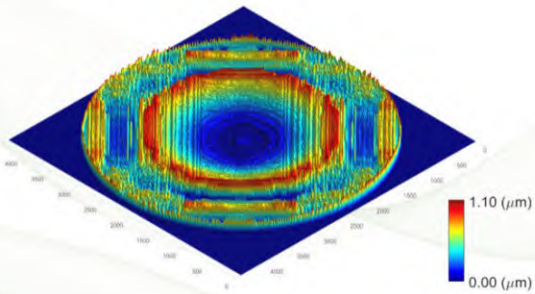


Learned Binocular-Encoding Optics for RGBD Imaging

Pipeline

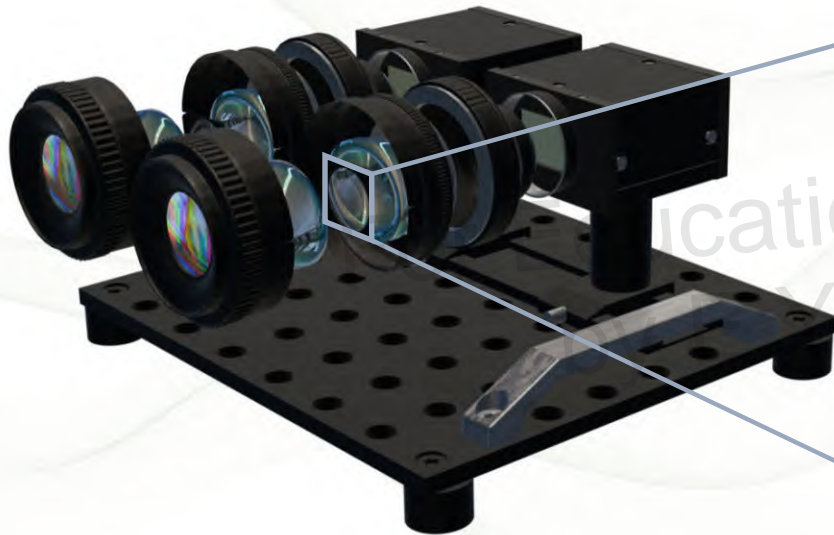
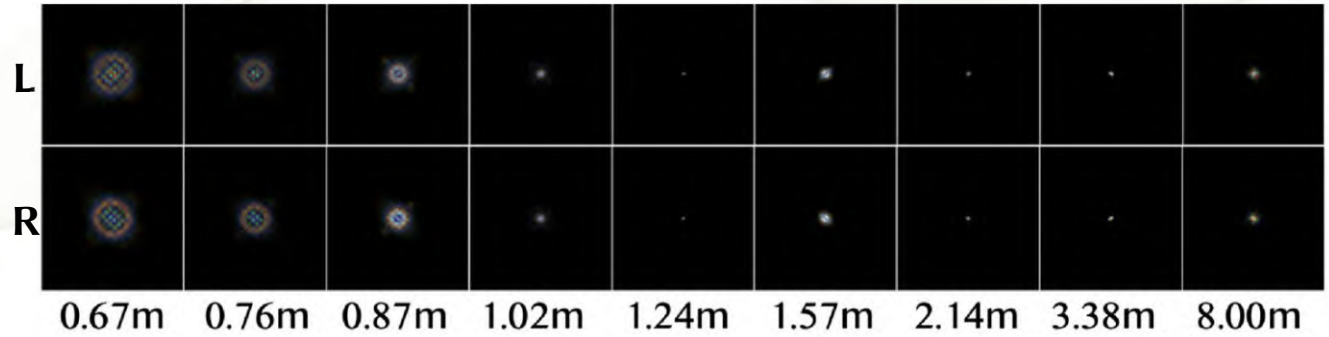


Customized Camera Prototype



3D Profile of Left DOE

PSFs

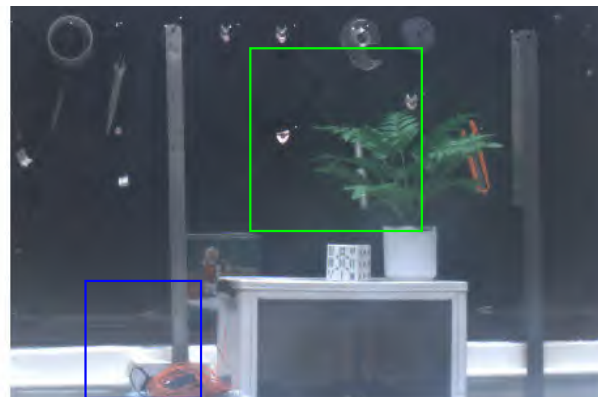


Experimental Results

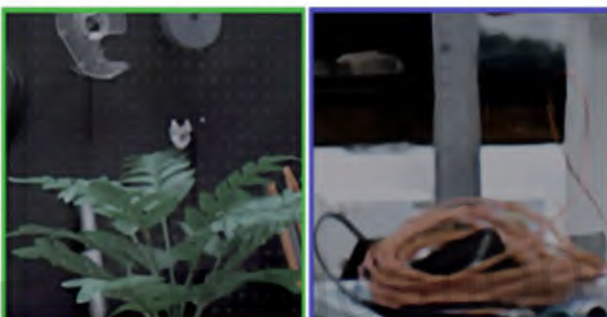
Conventional
Lens



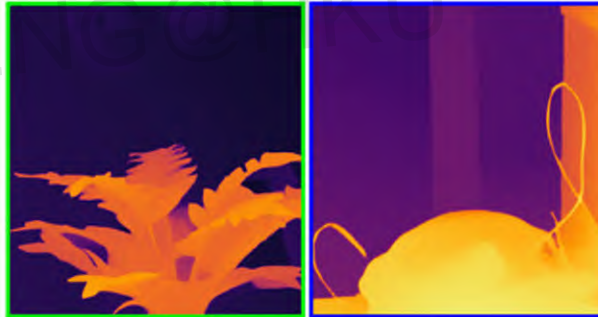
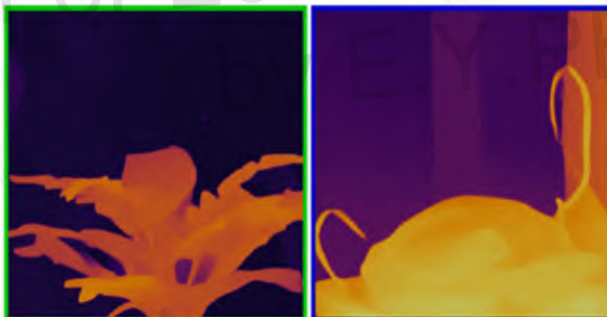
Lens w/
rank2-
coded DOE



AiF-RGB

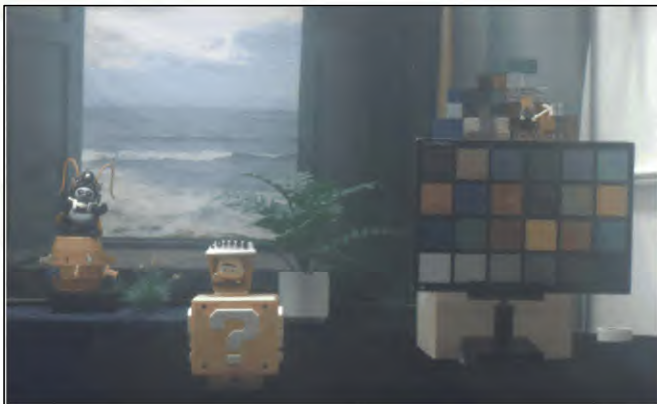


Depthmap

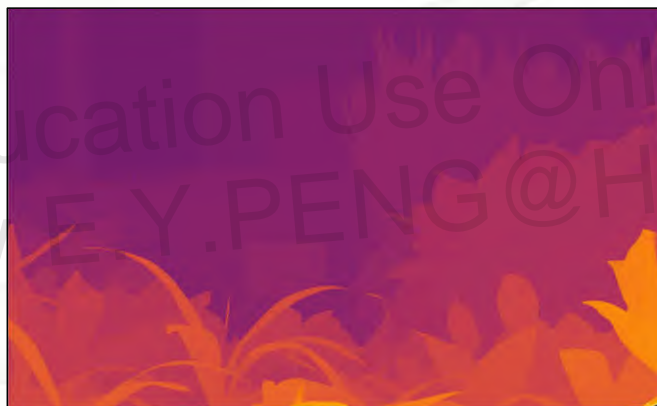


Experimental Results

→ Reconstructed Captured ←



AiF-RGB



High-fidelity Depth

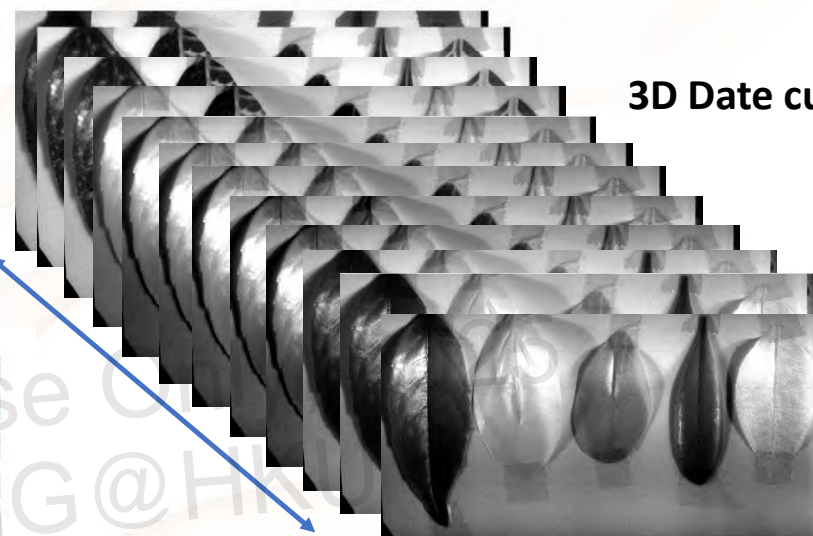
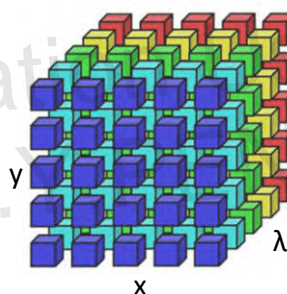
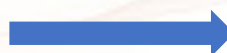
Why Snapshot Hyperspectral Imaging

Application scenario: Identify which leaves are fake.

Approach: Hyperspectral imaging.



RGB imaging



3D Data cube

Hyperspectral imaging

Why Snapshot Hyperspectral Imaging

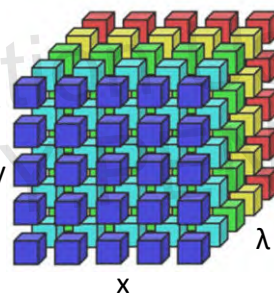
Application scenario:

Identify which leaves are fake.

Approach: Hyperspectral imaging.

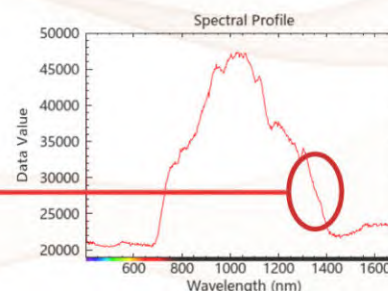


RGB imaging



True Fake Fake True Fake

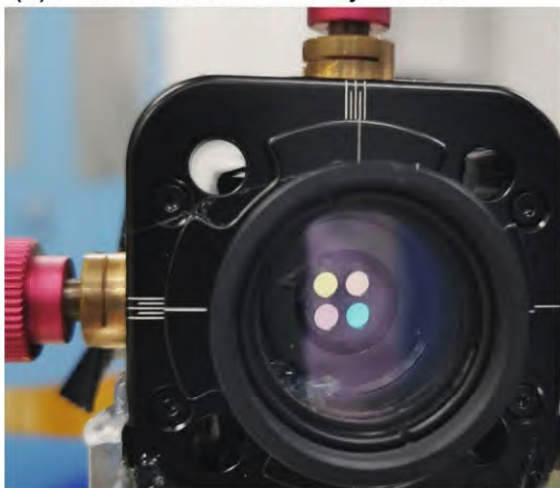
Hyperspectral imaging



Hyperspectral data cube (3D) beyond sensor capability (2D).

Learned Multi-aperture Color-coded Optics for Snapshot Hyperspectral Imaging

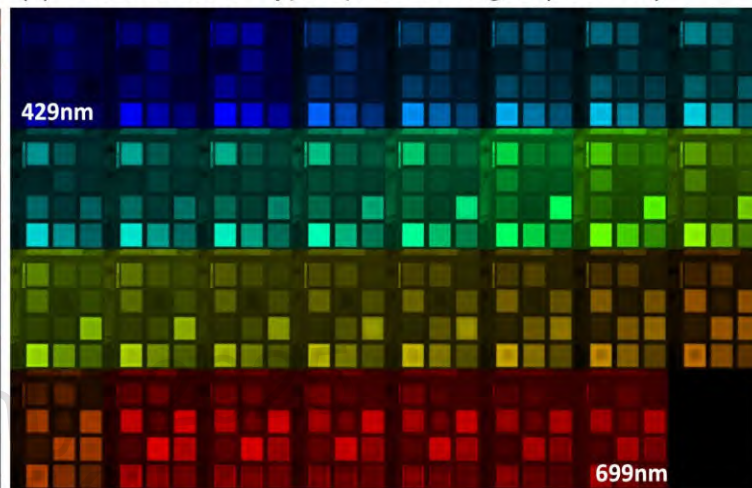
(a) Color Coded DOE Array Camera



(b) Reconstructed RGB Image



(c) Reconstructed Hyperspectral Images (zoom-in)



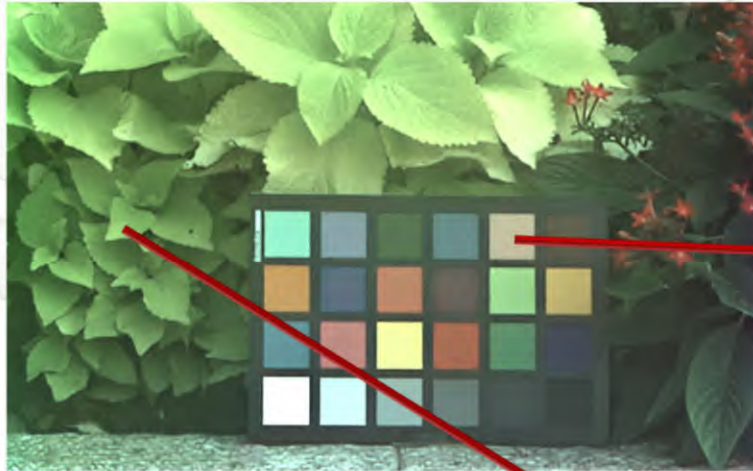
<https://whywww.github.io/ArrayHSI/>

Zheng Shi*, Xiong Dun*, Haoyu Wei, Siyu Dong, Zhanshan Wang,
Xinbin Cheng, Felix Heide, Yifan Peng, ACM TOG, 2024



RGB and Hyperspectral Imaging Results Highlight

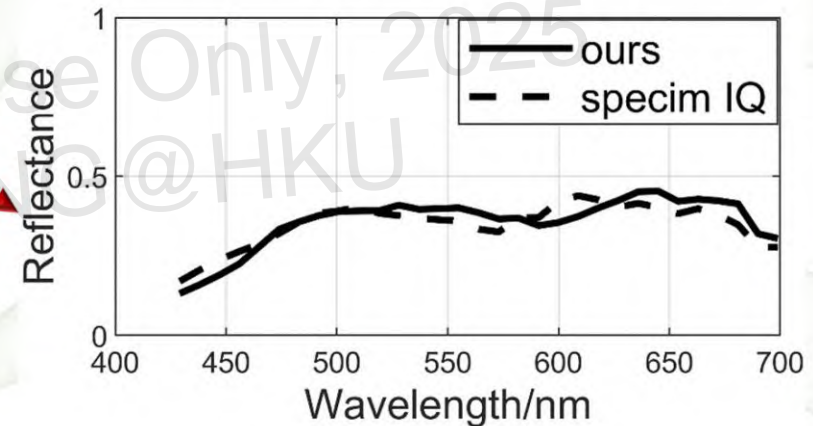
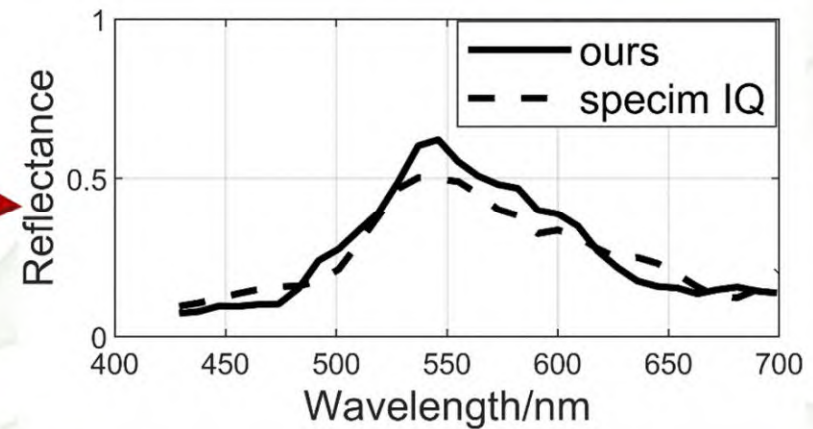
Ours



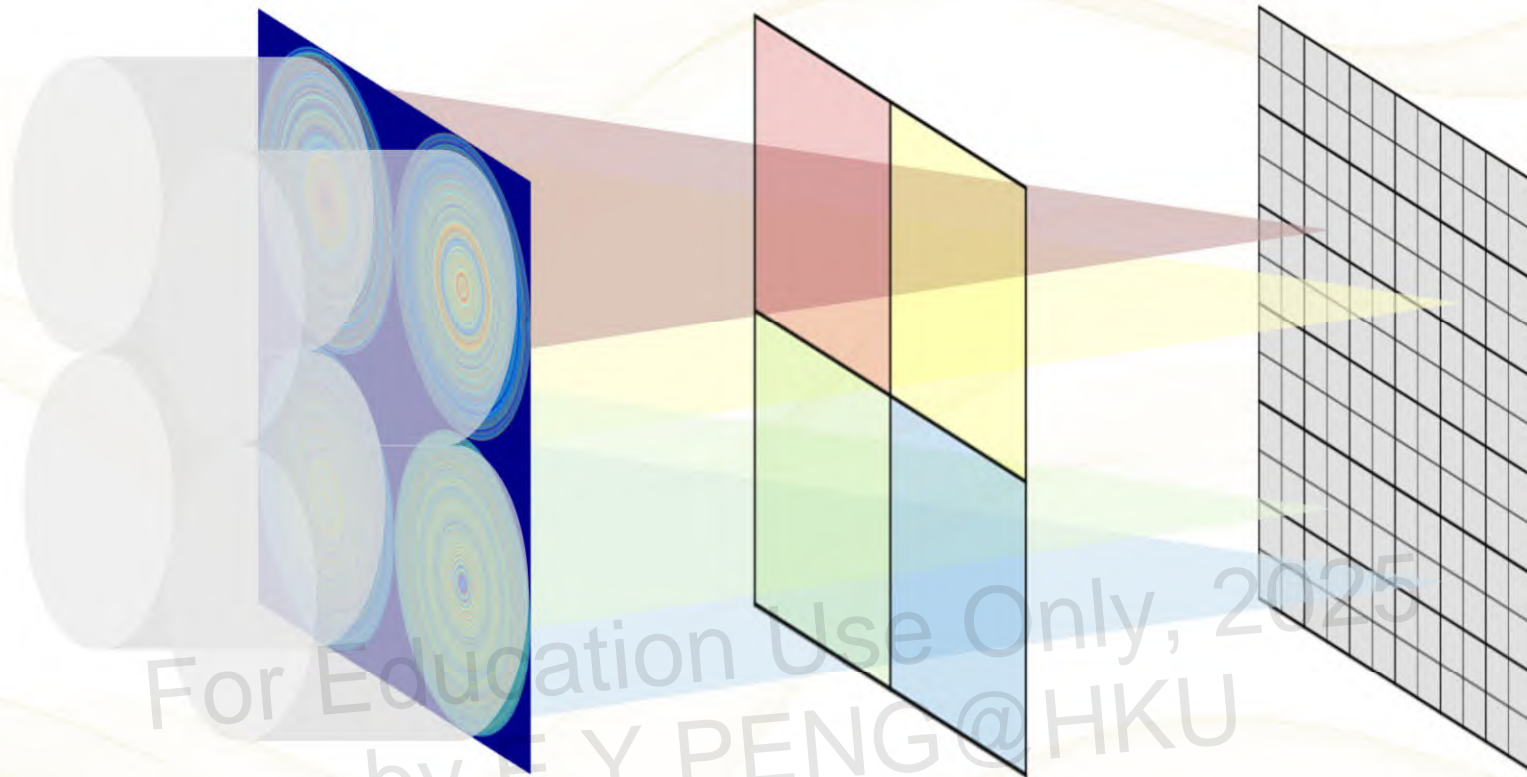
429 nm



Specim IQ



Tech Flow Overview

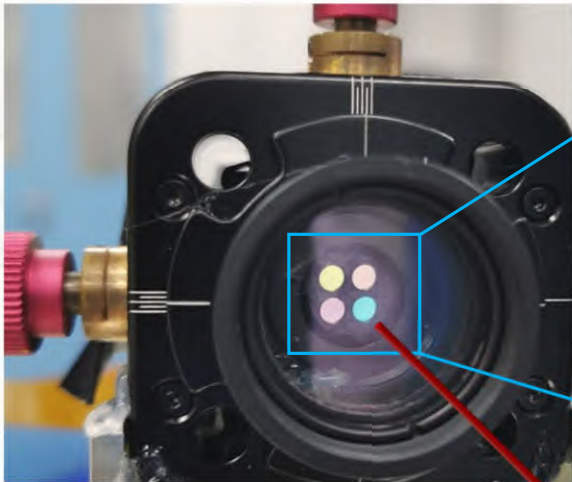


For Education Use Only, 2025
by E.Y.PENG@HKU

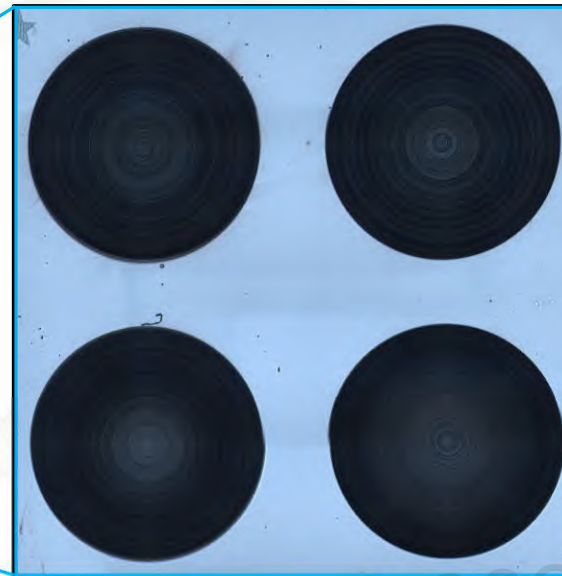
Spatial and spectral coding are fully **decoupled**

Aperture-wise filter also **avoids the high cost** of customized Bayer filter

Camera Prototype

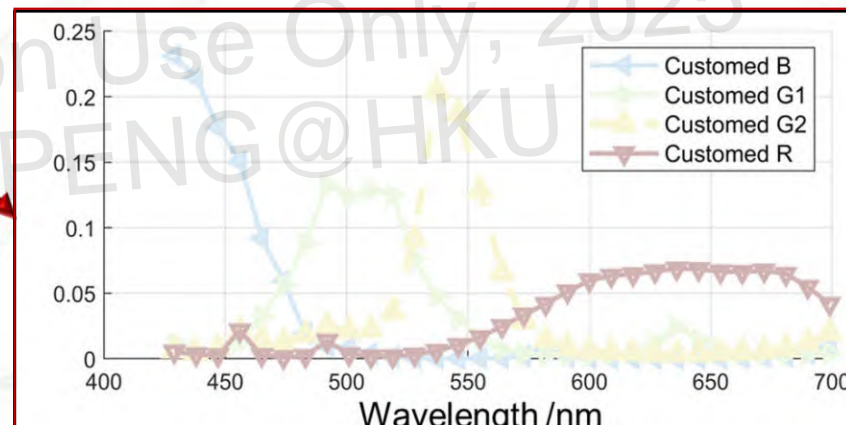


Aperture size: 3mm
EFL: 20mm
FOV: 20°

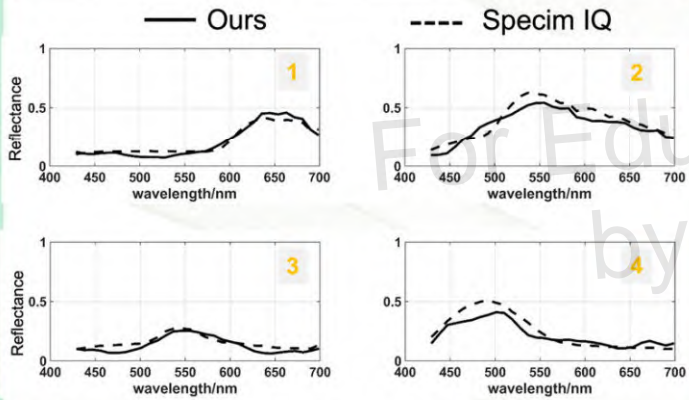
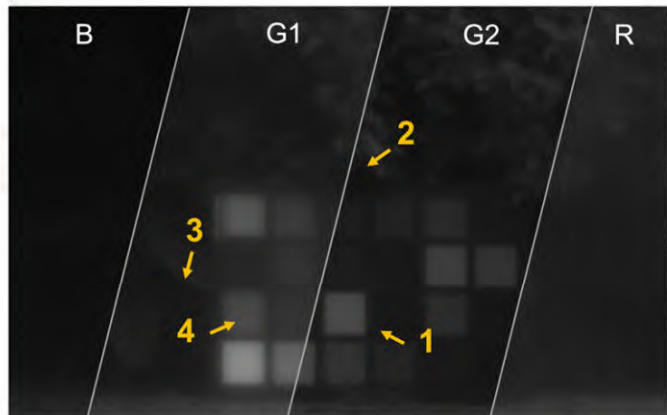


Microscope image of the DOE array

Customed spectral response curves



Experimental Results

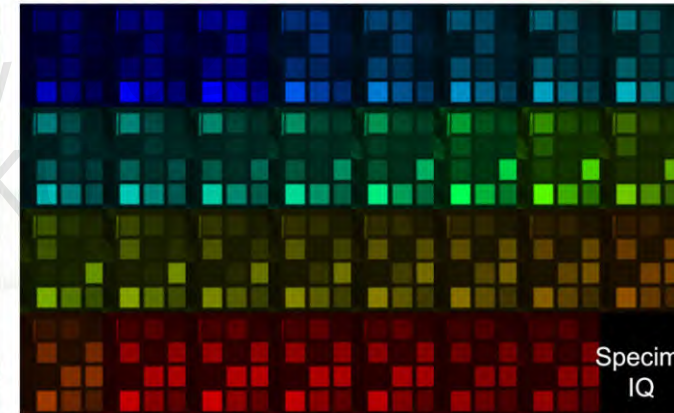
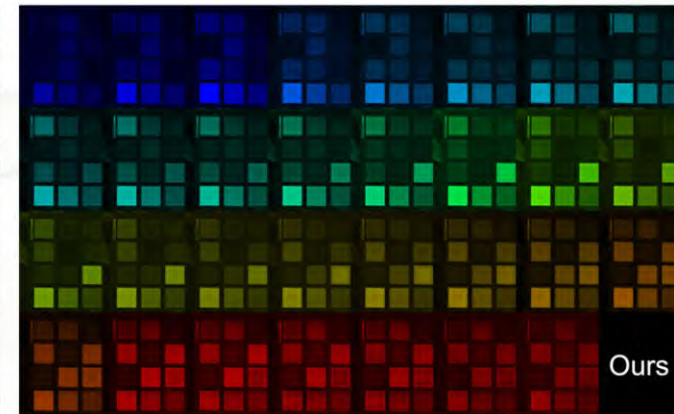


ACM TOG, 2024

RGB image

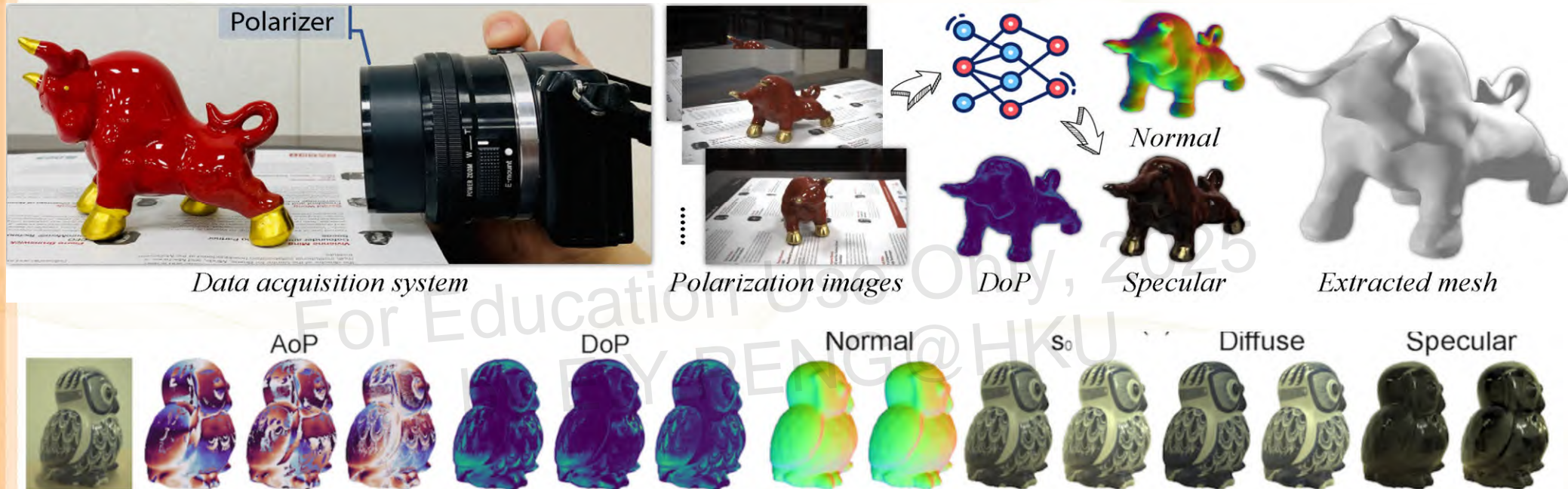


HSI images

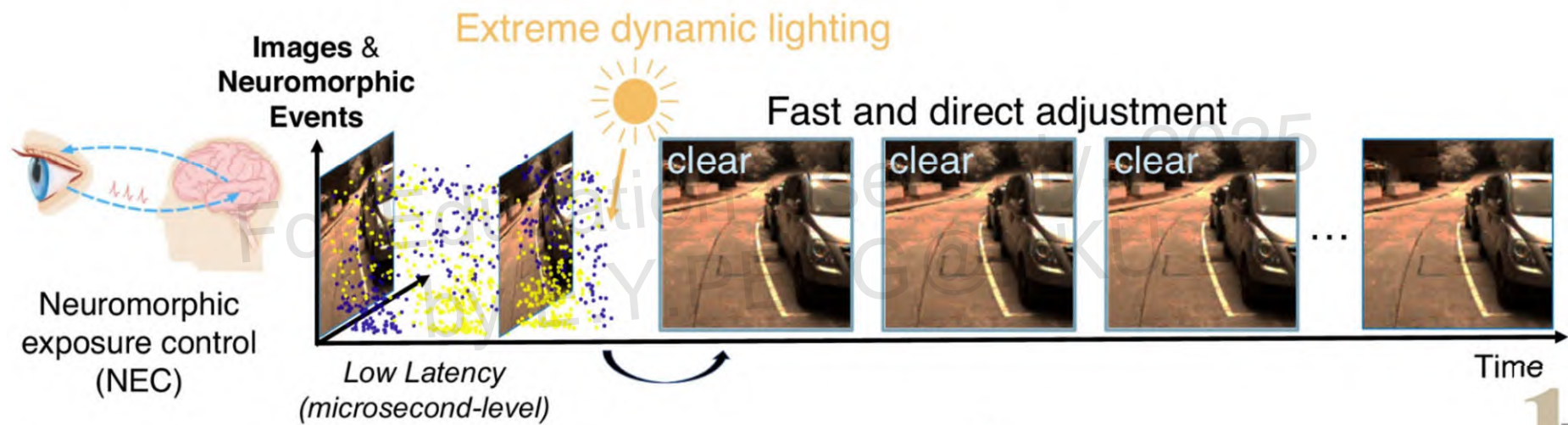
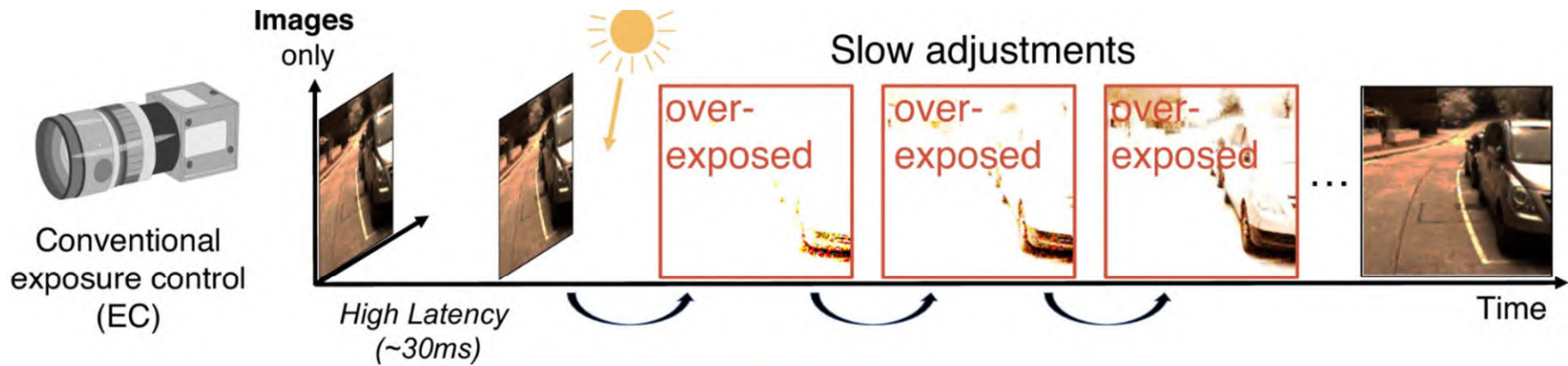


Glossy Object Reconstruction with Polarization Cues

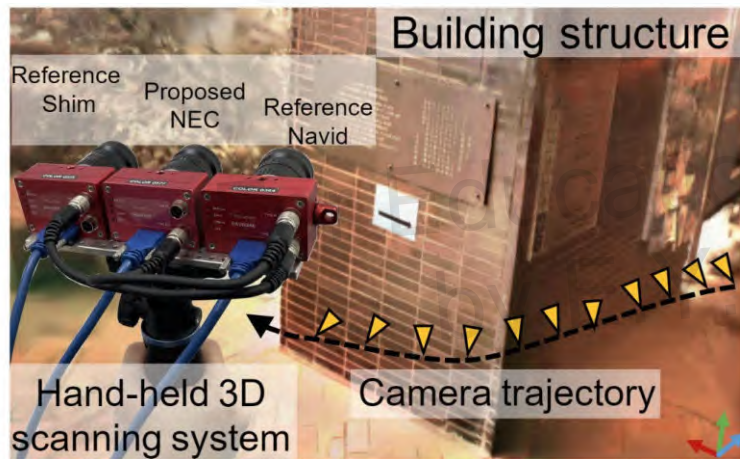
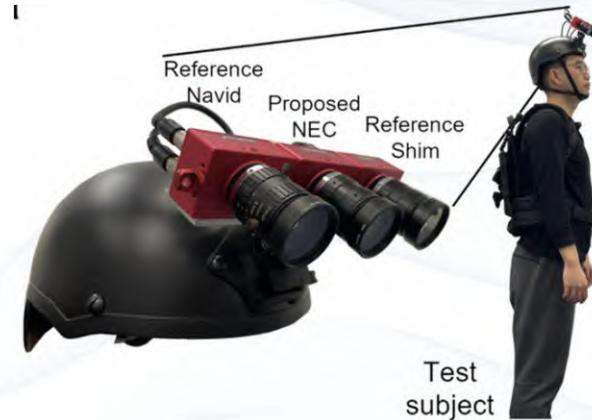
- **Cost-effective** polarized acquisition: RGB camera + single linear polarizer
- A single polarized image per view + NeRF + E2E polarization rendering = Accurate reconstruction & radiance decomposition



* Bojian Wu, Yifan Peng, Ruizhen Hu, and Xiaowei Zhou, IEEE CVPR, 2025 (Highlight)



Machine Vision in Extreme Bright: Embodied **Neuromorphic** Synergy



* Lin et al. Seeing in Extreme Bright: Embodied Neuromorphic Synergy for Lighting-robust Machine Vision, Nature Comms., 2024

Take-home Messages (**Imaging Beyond RGB**)

Why ?

- Thinner, lighter products
- Large, flexible design
- Optimal visual performance

What ?

- Possibly compact form factor optics
- High-definition image reconstruction
- Domain-specific** visual & sensory applications

How ?

- Wave propagation (DOE or Freeform lens)
- Auto-tuning & Data-driven advances
- Physics + AI **differentiable** optimization

Bunny
Lover!



WeChat Official Account: *IntelligentOptics*

Intelligent Optics Sharing



感知 | 计算



成像 | 显示



Thank
You

evanpeng@hku.hk

Thank
You





Department of Electrical and Electronic Engineering

電機電子工程系

WORKSHOP ON

FRONTIERS of IMAGE SCIENCE AND VISUAL COMPUTING 2025

15 – 16 April 2025 10:00 am

Multi-purpose Zone Room 1, 3/F, Main Library, The University of Hong Kong (HKU)

Chair: Prof. Evan Y. Peng @ HKU EEE

ACKNOWLEDGEMENT:



FRONTIERS OF IMAGE SCIENCE AND VISUAL COMPUTING 2025

15 – 16 April 2025 10:00 am

Multi-purpose Zone Room 1, 3/F, Main Library, The University of Hong Kong (HKU)

Chair: Prof. Evan Y. Peng @ HKU EEE x CS



FISVC2025

Details & Registration

www.eee.hku.hk/20250415

WORKSHOP OBJECTIVES

- Encourage Innovative Spirit
- Promote Excellence and Sustain Quality
- Strive for Improvement
- Connect Communities

Meet Our Speakers!



David FORSYTH
University of Illinois Urbana-Champaign (UIUC)
The Joy of Latents: What Image Generators "Know" and Don't "Know", and Why It Matters



Xiaojuan QI
The University of Hong Kong (HKU)
Learning 3D Representations from Videos



Hongzhi WU
Zhejiang University (ZJU)
Differentiable Acquisition of Appearance and Geometry



Taku KOMURA & Hongbo FU
HKU & HKUST
ACM SIGGRAPH Asia 2025 Pitching



Wenzheng CHEN
Peking University (PKU)
Open-World Rain Simulator

He SUN

Peking University (PKU)

Diffusion Posterior Sampling for Computational Imaging with Inaccurate Priors or Physics



Yinqiang ZHENG

The University of Tokyo (UTokyo)
Security Concerns on Visual AI beyond RGB Domain



Tianfan XUE

The Chinese University of Hong Kong (CUHK)
Computational Photography in the Age of Foundation and Generative Models



Yuanmu YANG

Tsinghua University (THU)
Flat-Optics for Imaging with Extended Information



Ping TAN

Hong Kong University of Science and Technology (HKUST)
Towards Better Gaussian Splatting: Enhancing Geometry and Volumetric Rendering



Seung-Hwan BAEK

Pohang University of Science and Technology (POSTECH)
Computational Imaging for Graphics, Vision, Robotics, and Science



• Closing Remarks: **Prof. Kaibin HUANG**, Head of HKU-EEE

GENERATIVE

RENAISSANCE

ASIA.SIGGRAPH.ORG/2025



SIGGRAPH 香港
ASIA 2025
HONG KONG 香港

Conference 15 – 18 December 2025
Exhibition 16 – 18 December 2025
Venue Hong Kong Convention
and Exhibition Centre

Use Only, 2025
Shaping the future

Sponsored by



Organized by

