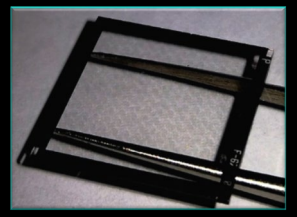
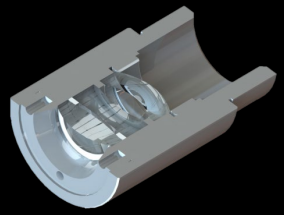
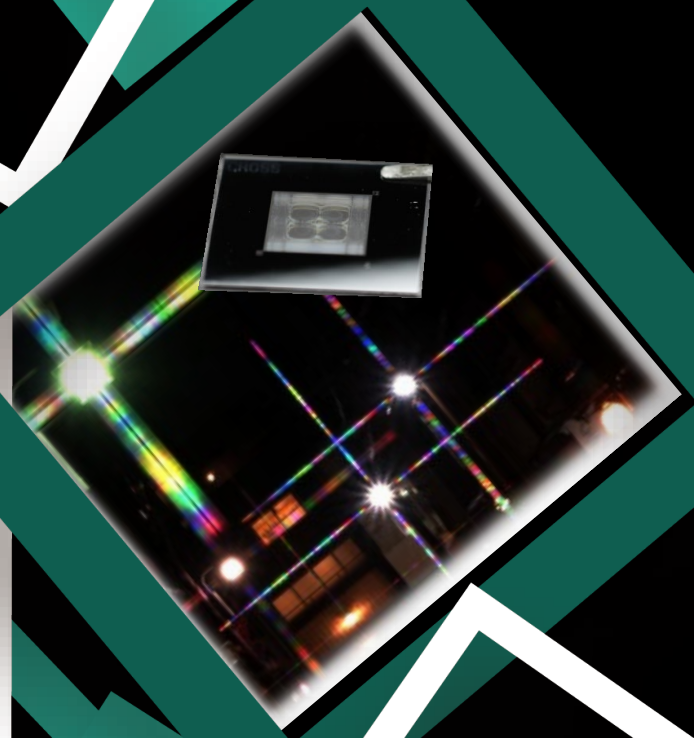









Computational Imaging

Lecture 24: End-to-end Computational Camera Design

点昀 POINT SPREAD

Qilin Sun（孙启霖）

香港中文大学（深圳）

点昀技术（Point Spread Technology）

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

- Step One: Joint Optimization of Optics and Image Recovery
  - Learned Large Field-of-View Imaging with Thin-Plate Optics
- Step Two: Compressive and Optically Coded Super-resolution SPAD Camera
  - Compressive and Optically Coded Super-resolution SPAD Camera
- Step Three: Enabling Comprimization for Color Channels
  - Learning Rank-1 Diffractive Optics for Single-shot High Dynamic Range Imaging
- Step Four: End-to-end Complex Lens Design
  - End-to-end Complex Lens Design with Differentiable Ray-Tracing

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
点昀 POINT SPREAD

Step One: Joint Optimization of Optics and Image Recovery

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

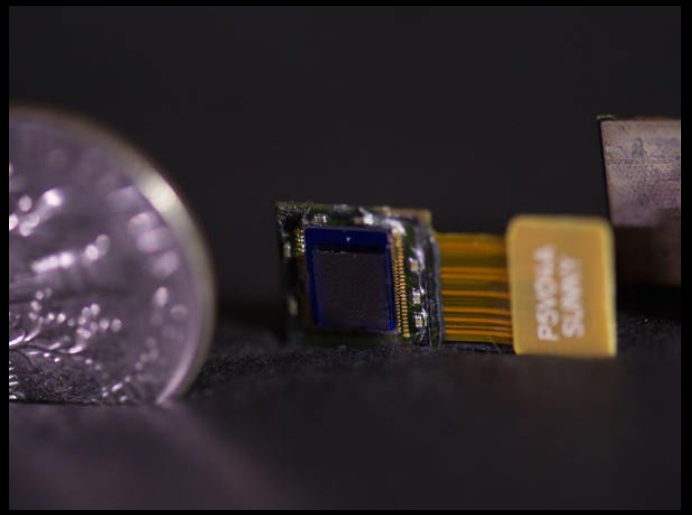
Thin-Plate Optics

- Diffractive optics
- Metasurfaces

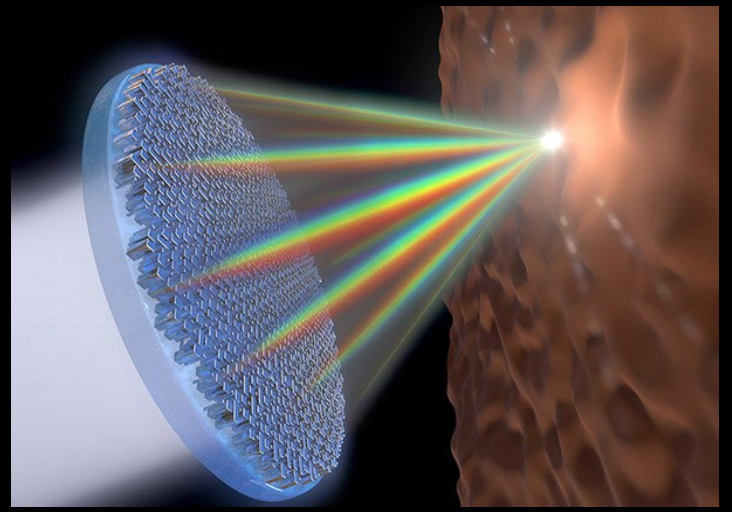
- Thin form factor
- Large design freedom

- Limited field-of-view
- Limited color fidelity
- Limited lens size

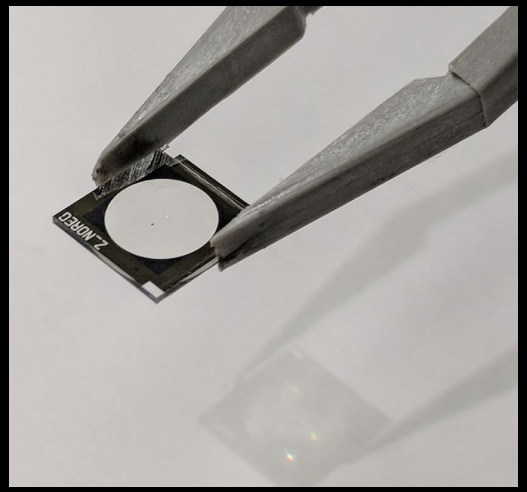
This work



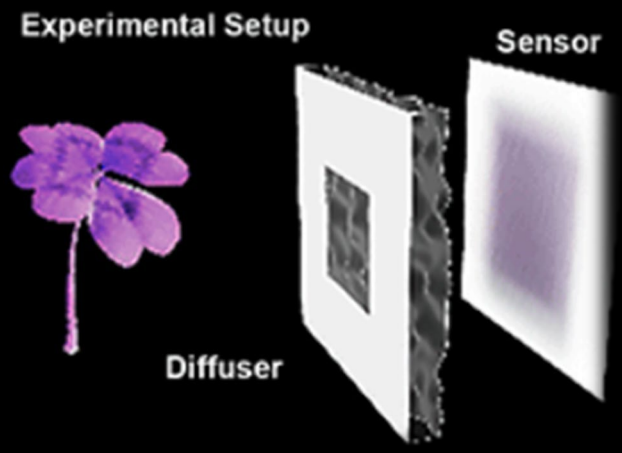
[Ayremlou et al. 2016]



[Capasso et al. 2017]



[Peng et al. 2016]
[Sitzmann et al. 2018]



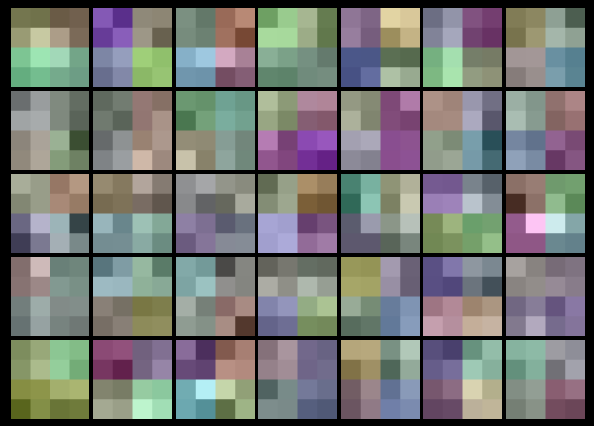
[Antipa et al. 2018]

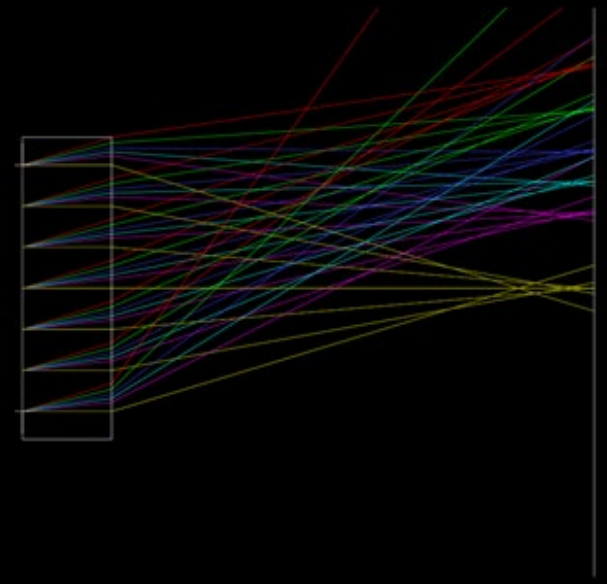
Wide Field-of-View Imaging

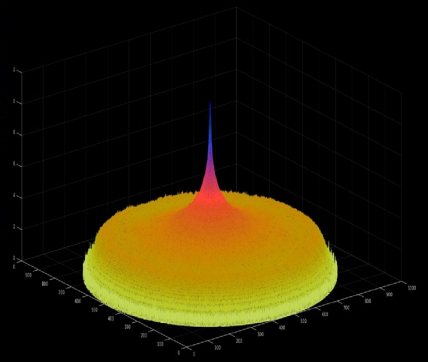
[Peng et al. 2016]
[Heide et al. 2016]
[Ayremlou et al. 2016]
[Sitzmann et al. 2018]

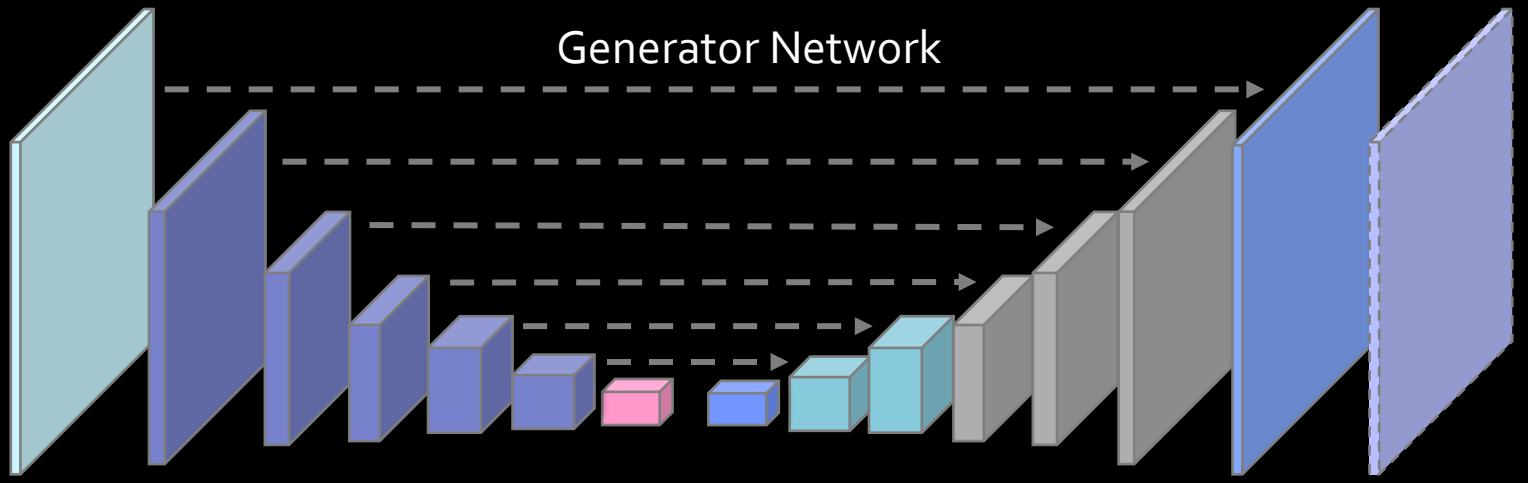
Co-design of Optics and Learned Recovery

Co-design of Optics and Learned Recovery

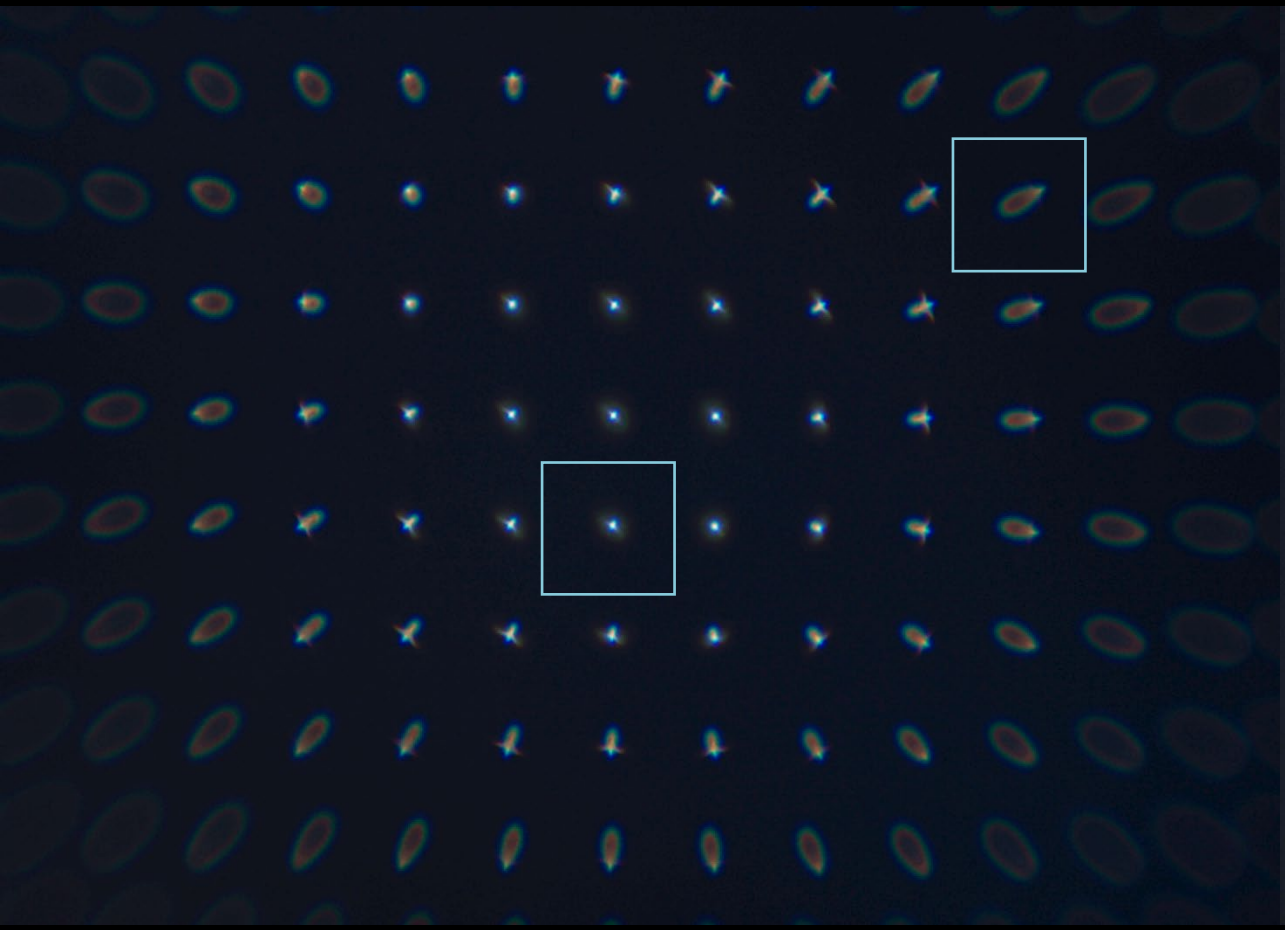


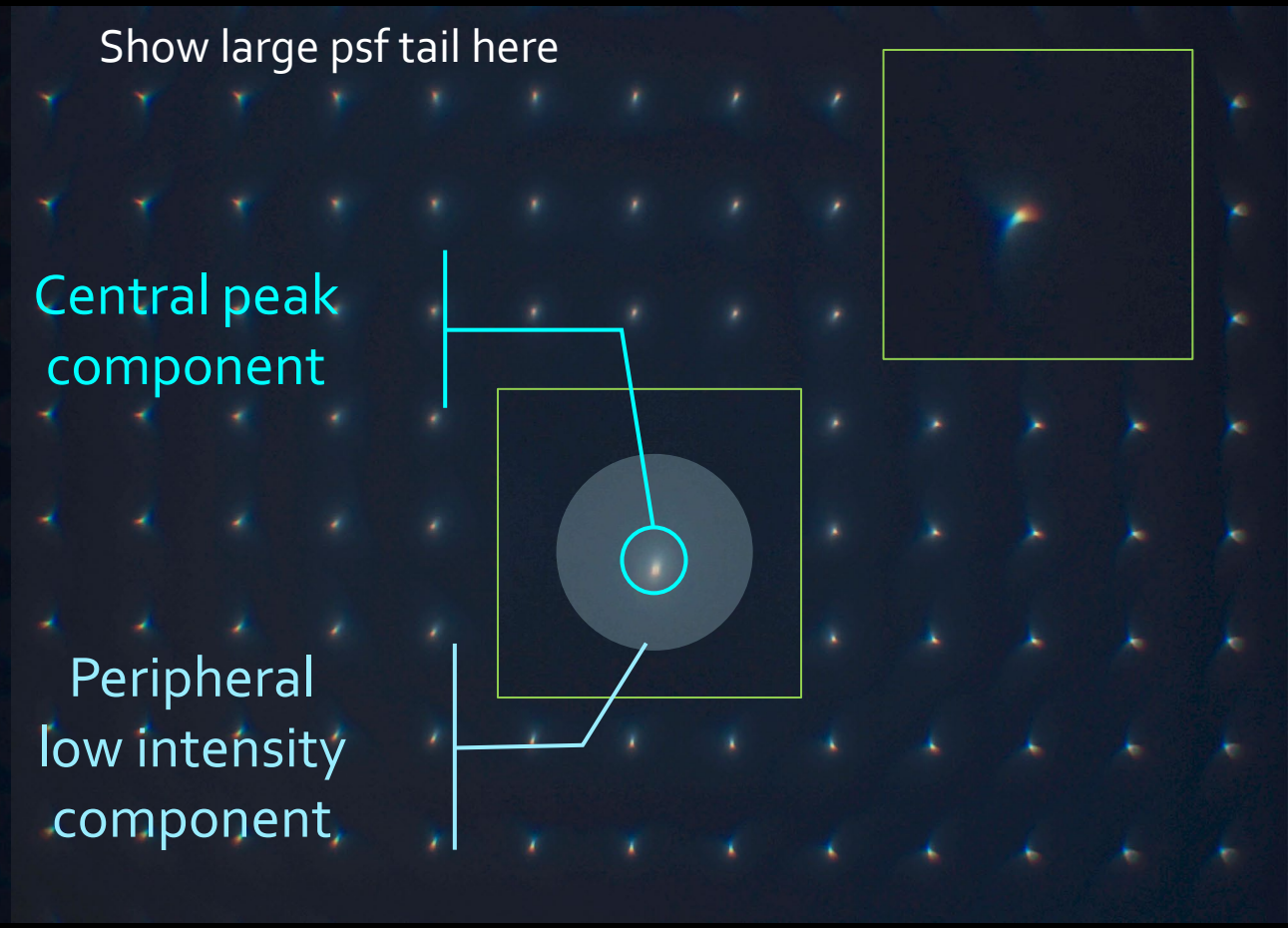






Angularly-Invariant PSF

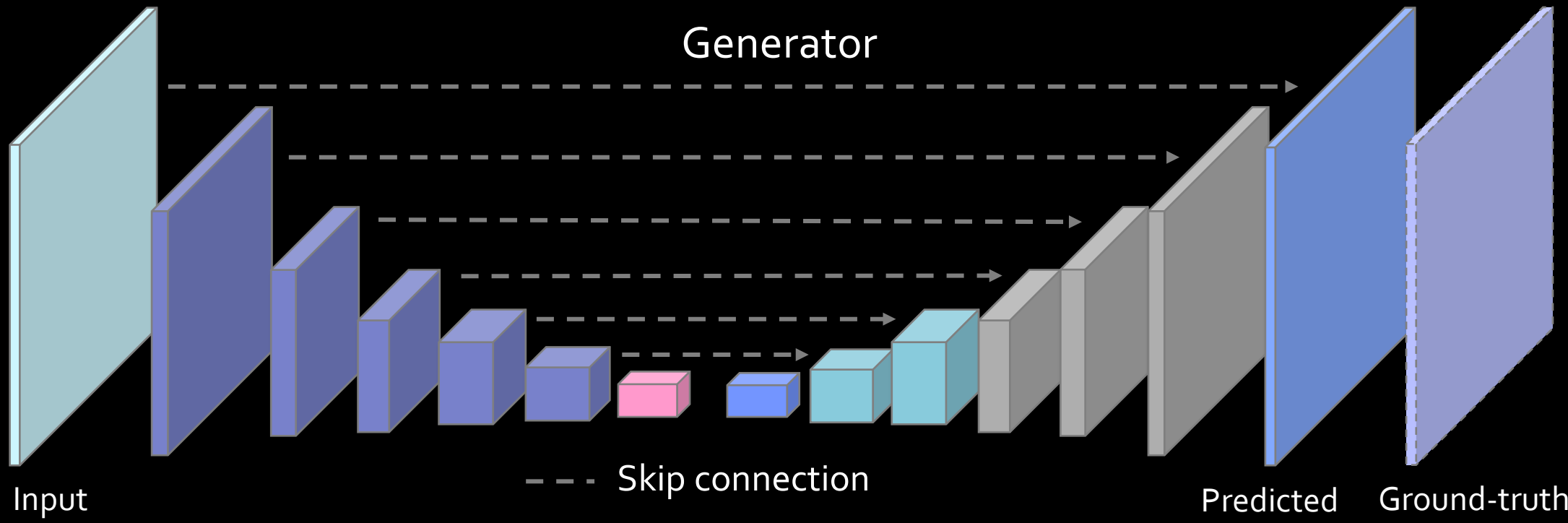


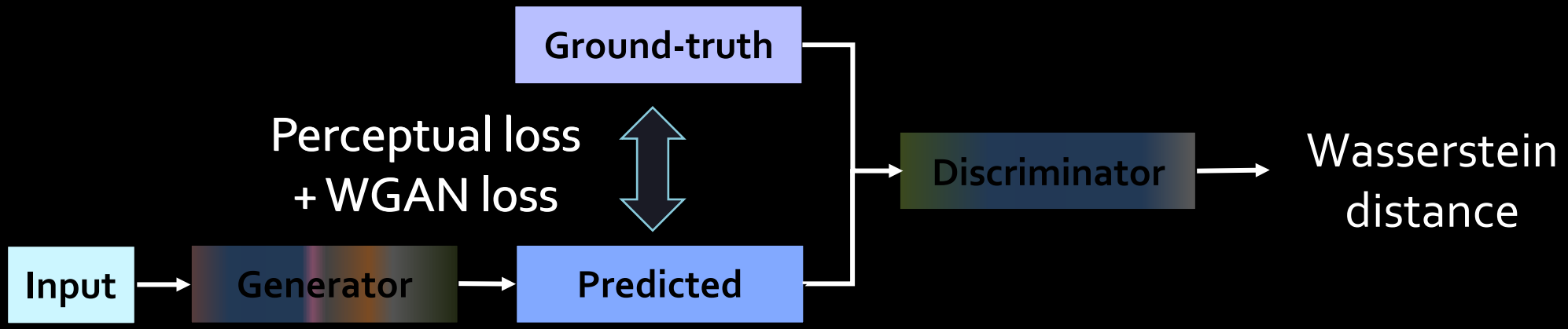


Aspherical lens (f = 50 mm)

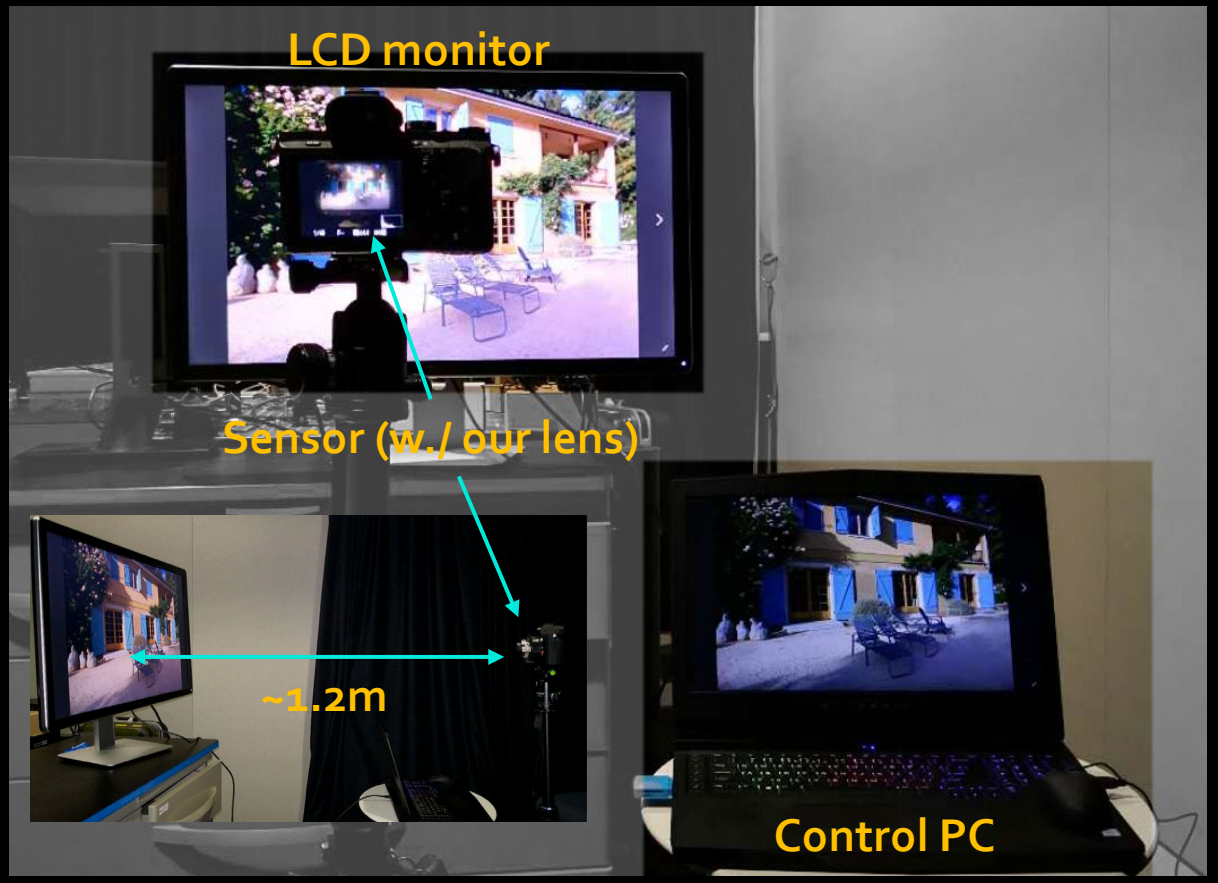
Our design lens (f = 43 mm)

Learned Reconstruction Framework



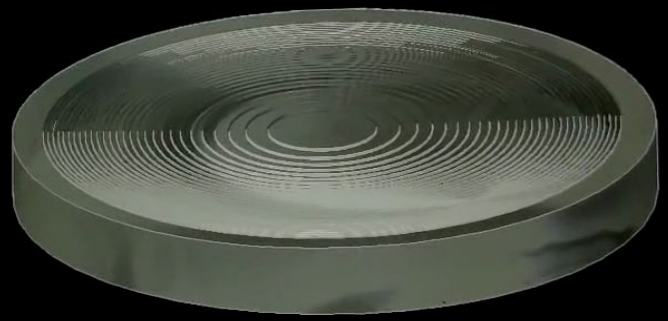


Automated Training Data Acquisition

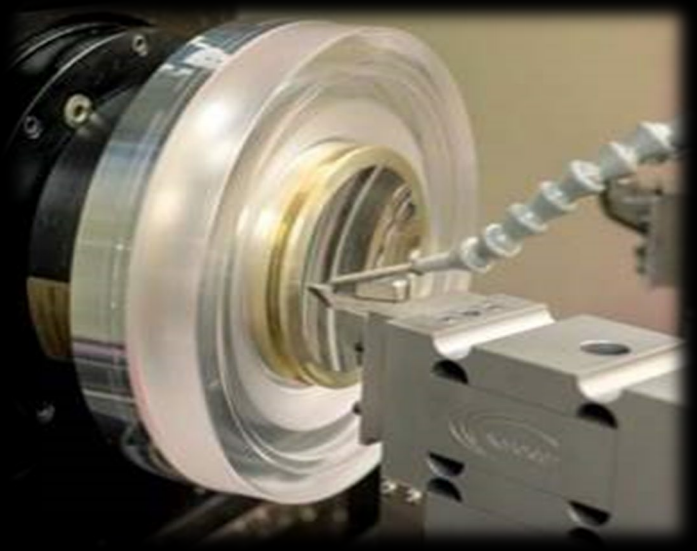


814 image pairs in total

Prototypes



- Clear aperture D : 23.4 mm
- Focal length f : 43.0 mm
- Field of view : 50+ deg.



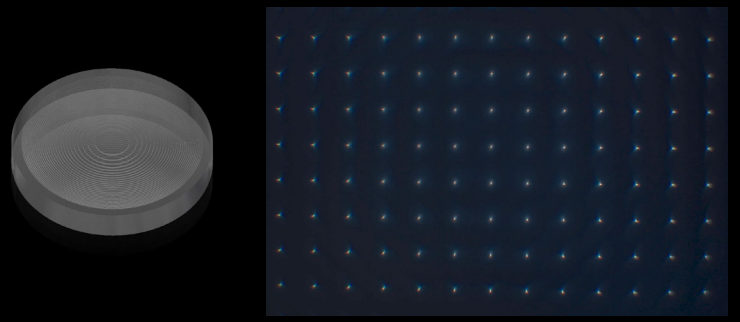
10 mm substrate　　3 mm substrate

Remarks

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

点昀 POINT SPREAD

- Large field-of-view + Single element



- Display-capture lab training

- High-quality color imaging

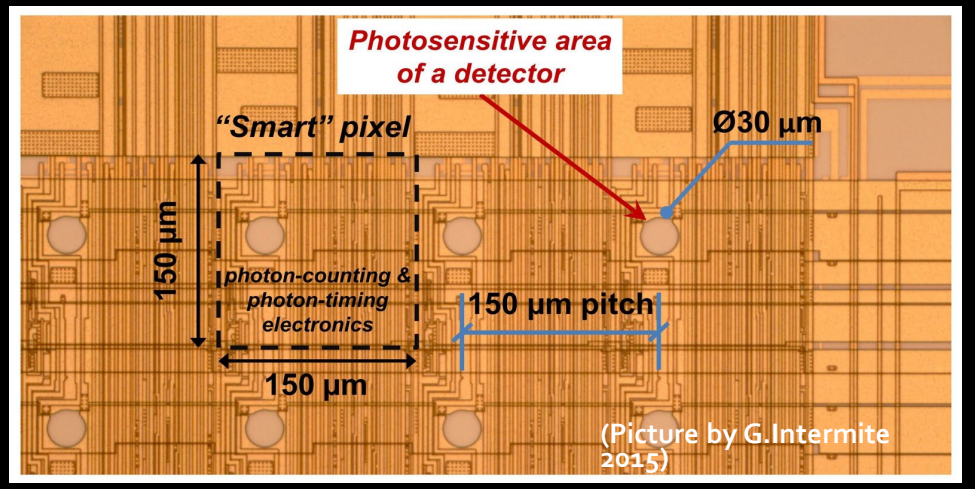
Step Two: The First End-to-end Optics for Single Wavelength

Compressive and Optically Coded Super-resolution SPAD Camera

Motivation

SPAD array sensor

(Picture from MPD)



Low fill-factor and LR image

=

Sampling HR image with low ratio

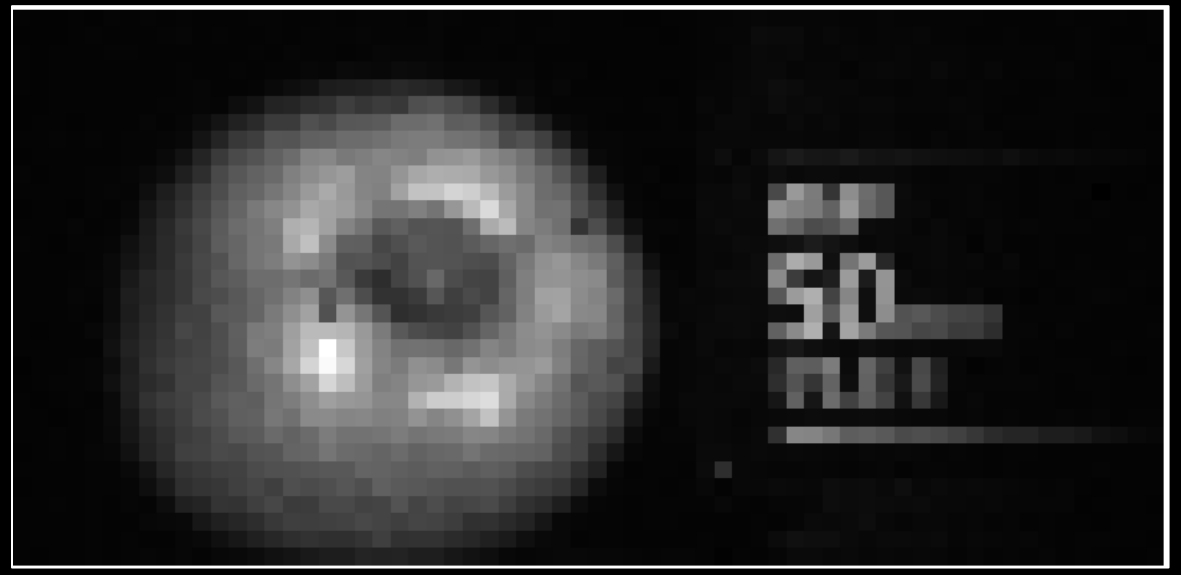
- Single photon level sensitivity
- Pico-second level time resolution
- Low spatial resolution
- Low fill-factor

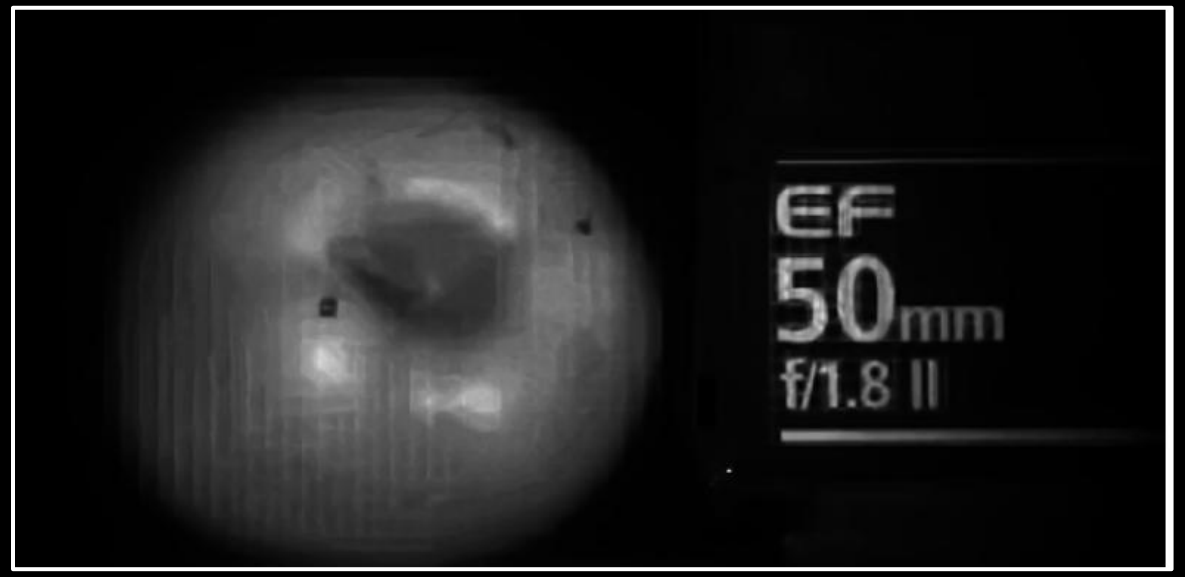
Optimal optics PSF encoding

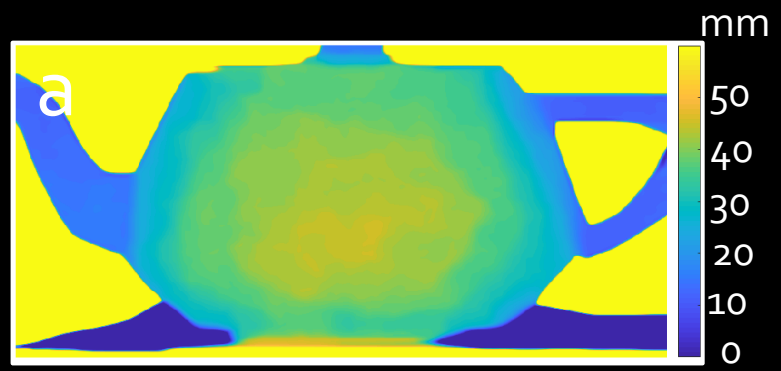
Deep Network based SR method

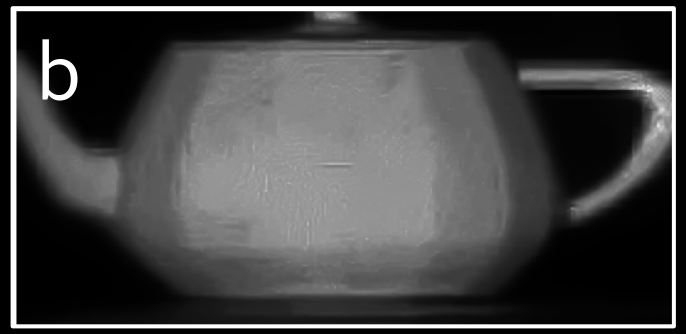
⊕ → Recover HR image with single-shot

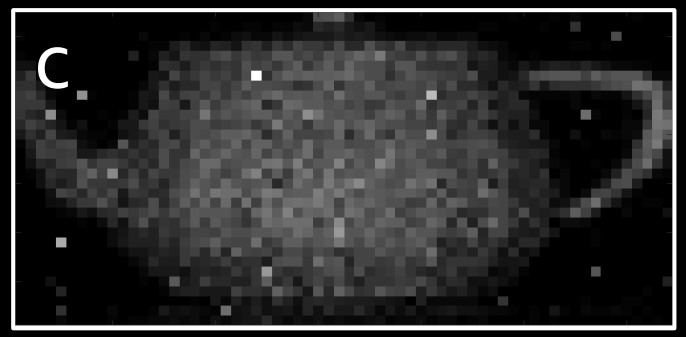
Compressive SPAD Camera—Results



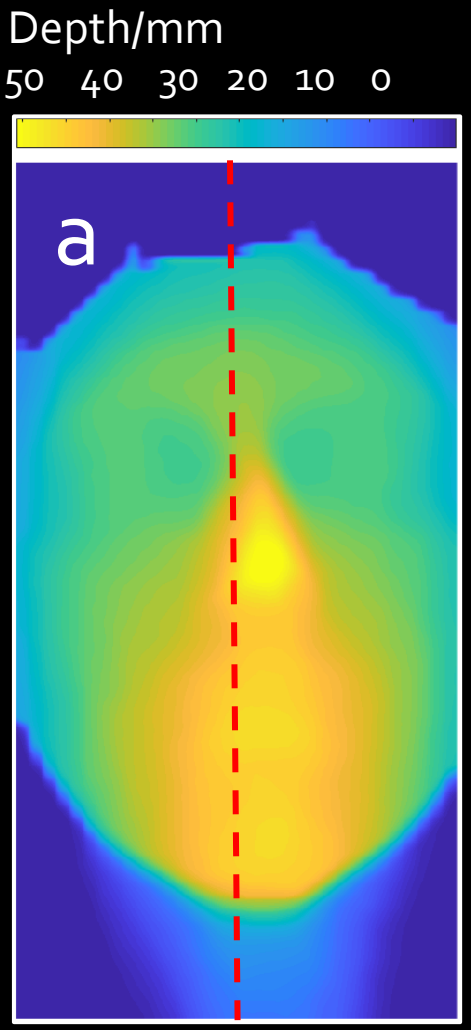


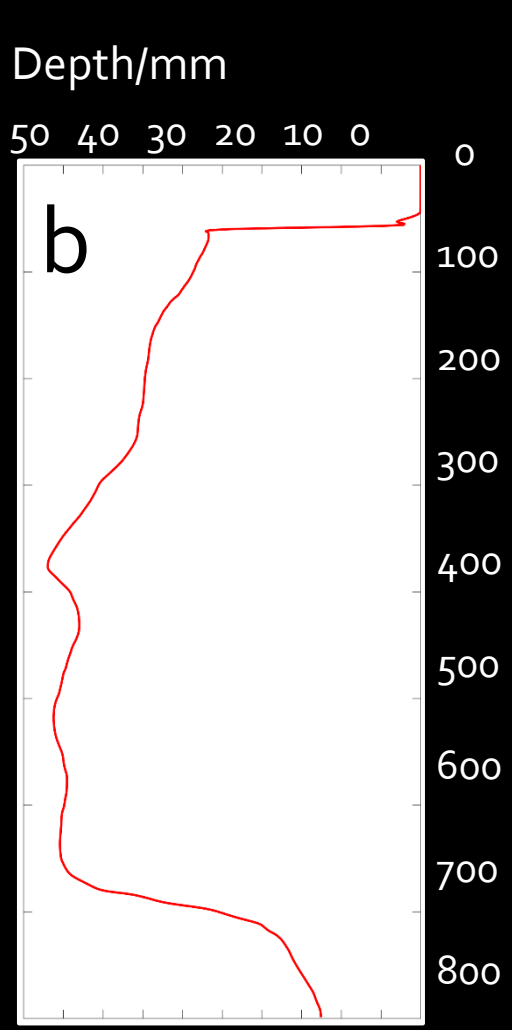


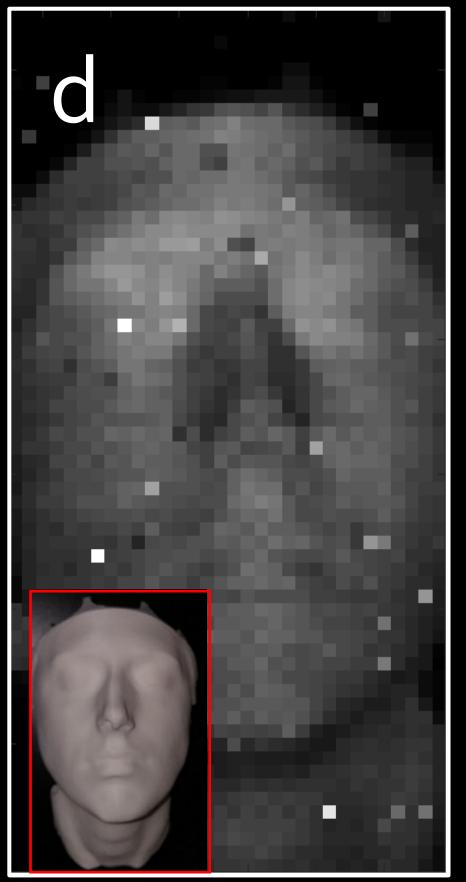




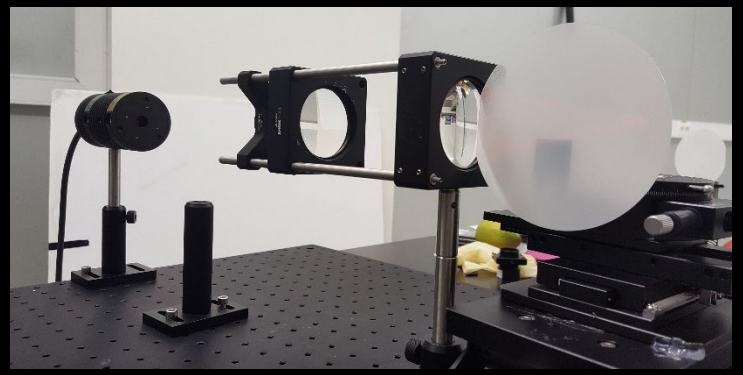
Compressive SPAD Camera—Results



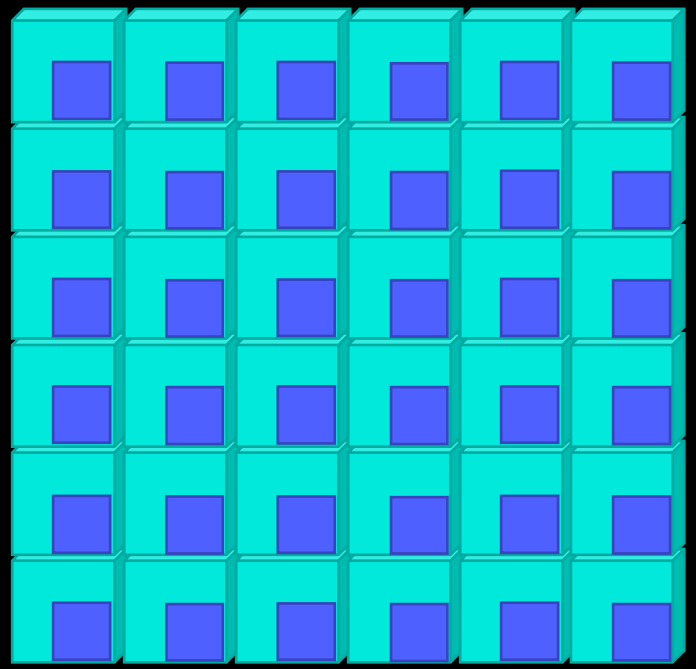




Compressive SPAD Camera—Results

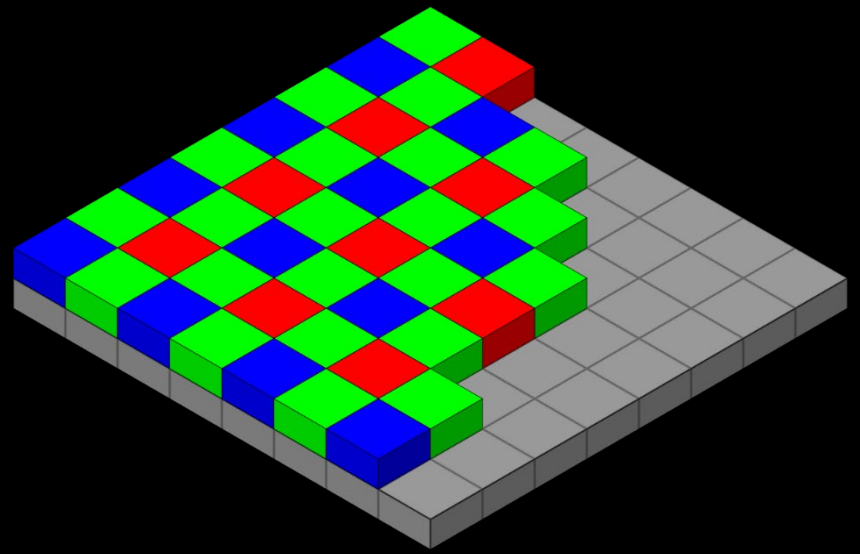


Sampling with a Physical Sensor Structure



Low fill-factor

- Aliasing artifacts
- Low efficient



Bayer filter

- Moiré pattern
- Mosaicing

Optimal Sampling for

Demosaicing

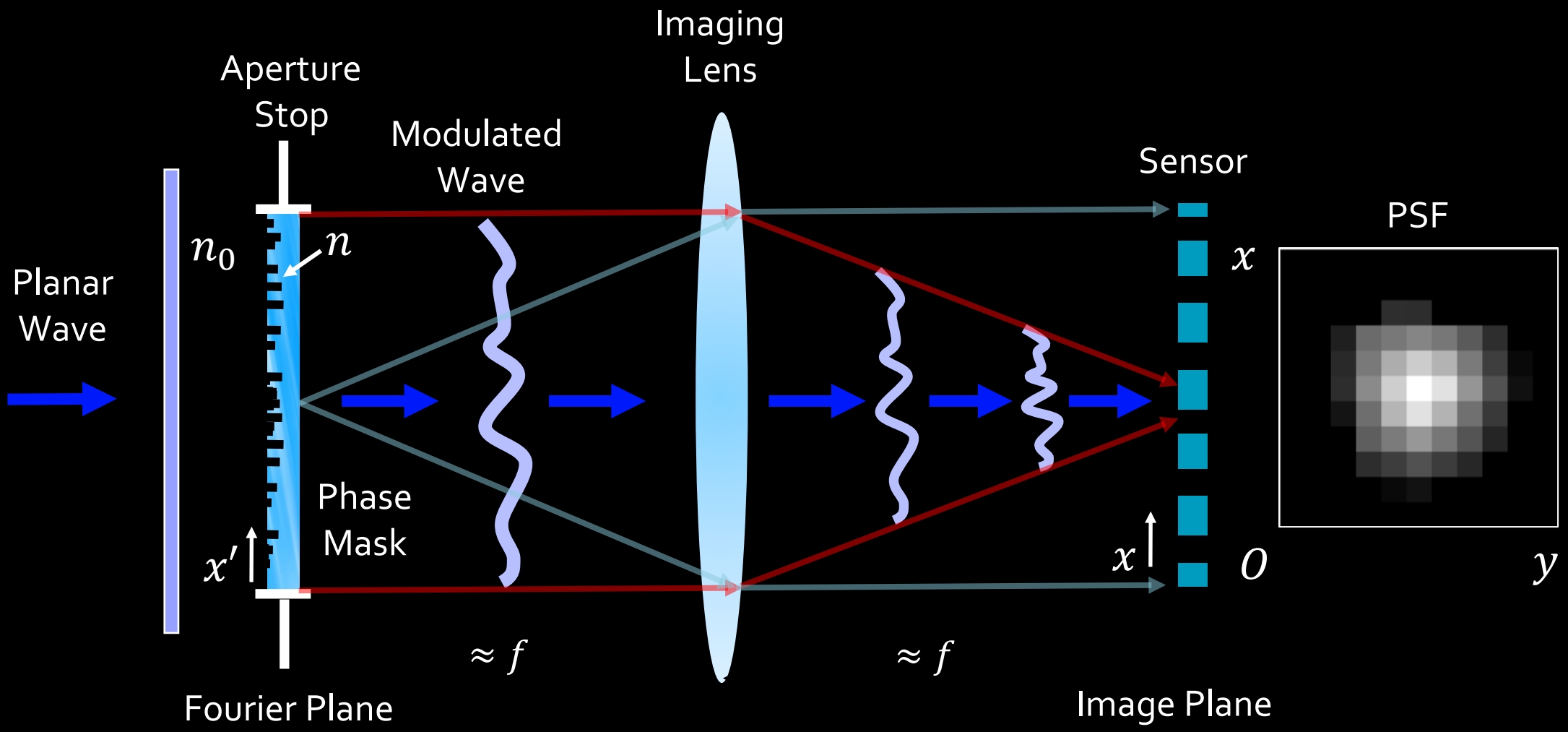
Anti-aliasing

Super-resolution

⋮

Etc.

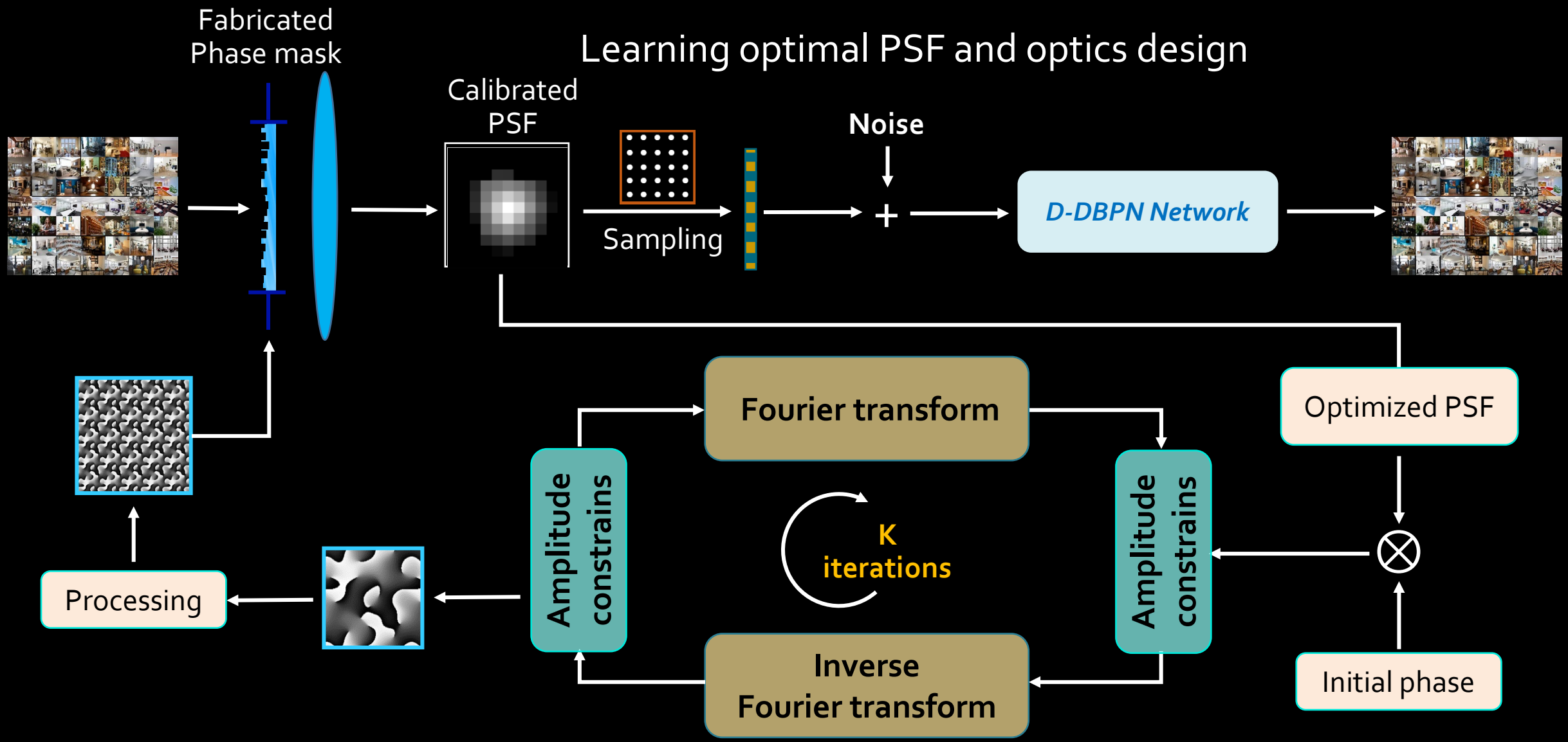
Optical Model



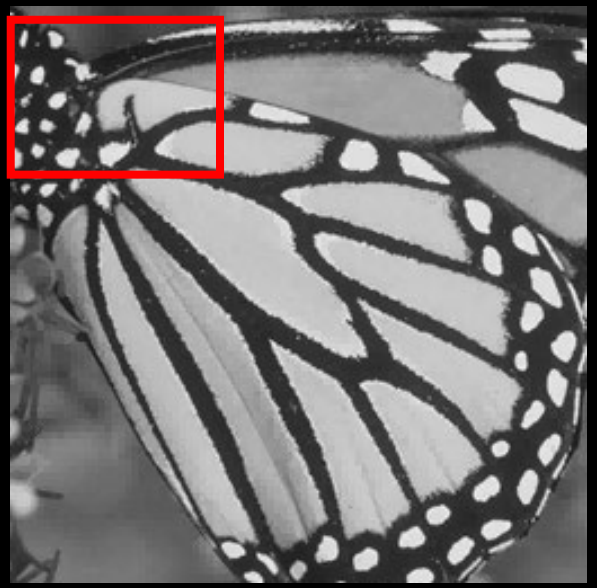
$\phi(x', y')$

$\rho_\lambda(x, y) \propto \|\mathcal{F}\{\phi(x', y')\}\|^2$

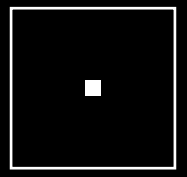
End-to-end Framework of SR SPAD Camera



Effectiveness of the Encoding Phase Mask



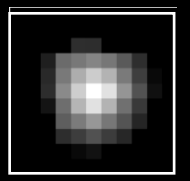
model	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSDS100 PSNR/SSIM
Low fill-factor	27.17/ 0.9019	23.97/ 0.8066	23.82/ 0.7691
Full fill-factor	29.77/ 0.9317	26.13/ 0.8442	25.59/ 0.8069
Ours	30.76/ 0.9399	26.91/ 0.8557	26.23/ 0.8198



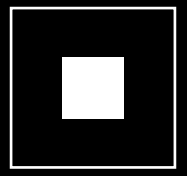
Low fill-factor

Low fill-factor (no mask)

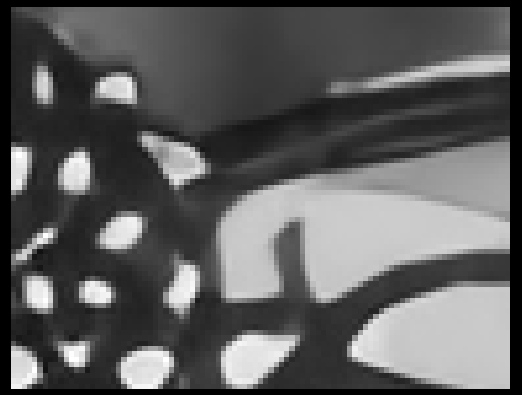
With phase mask



Ours



Full fill-factor



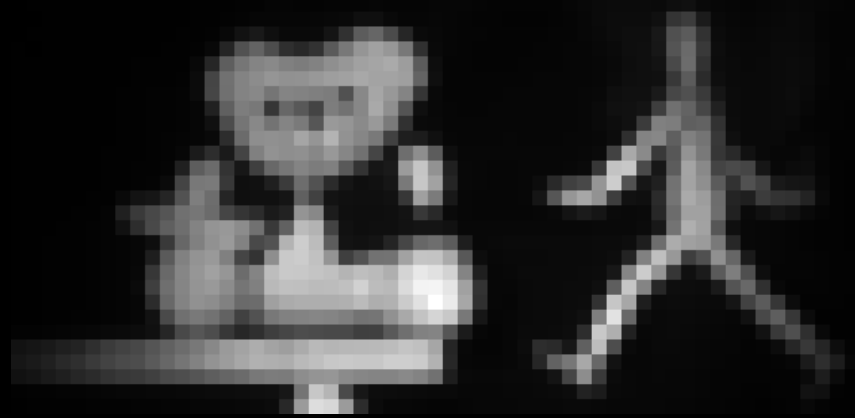
Full fill-factor

Ground Truth

Effectiveness of the Encoding Phase Mask

Raw image without mask

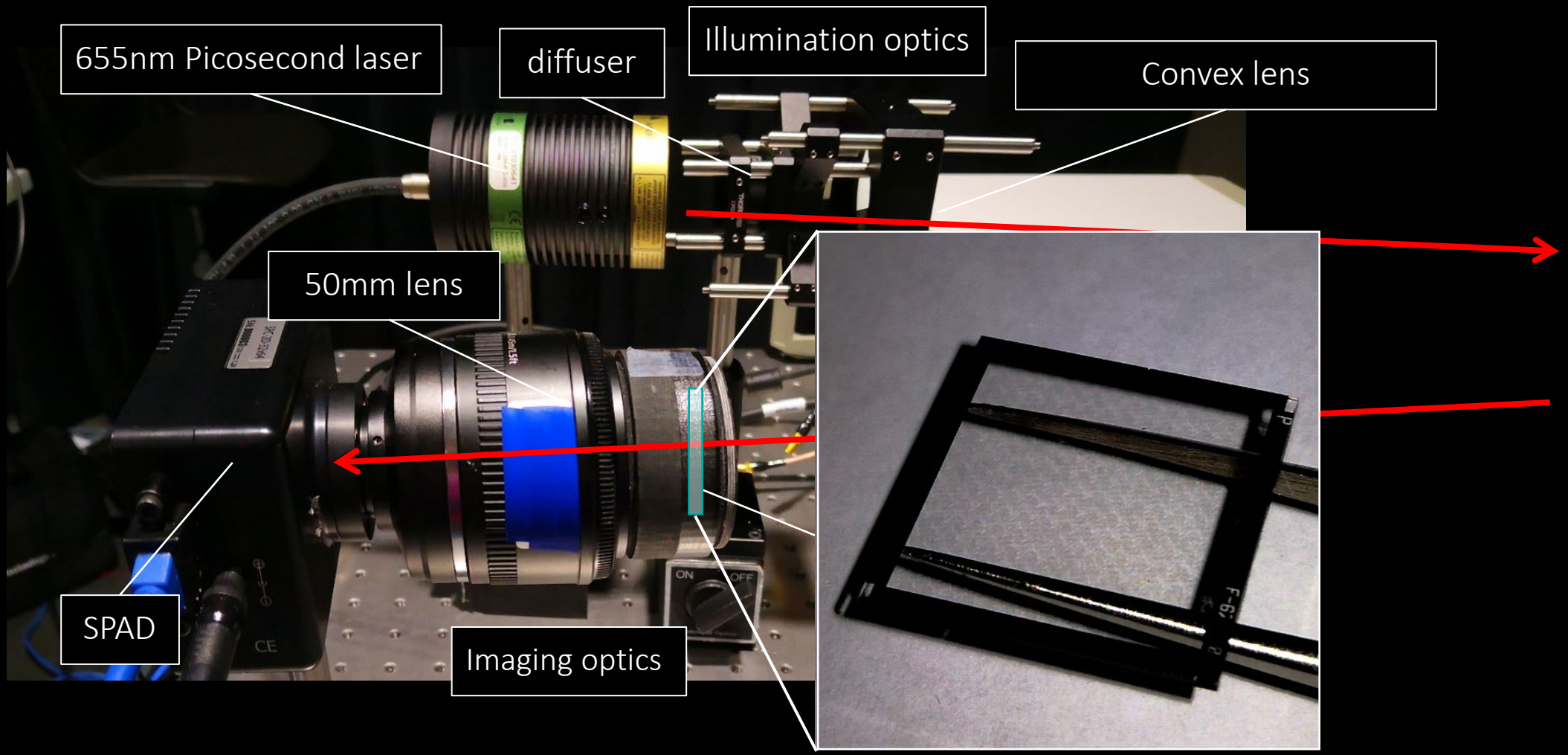
Result without mask



Raw image with mask

Result with mask

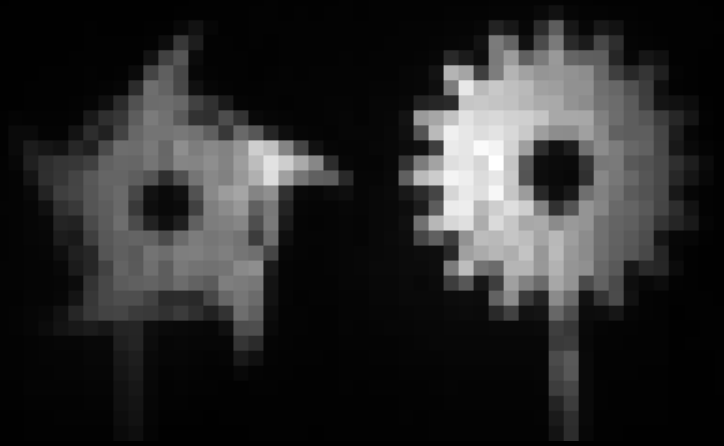
Prototype System



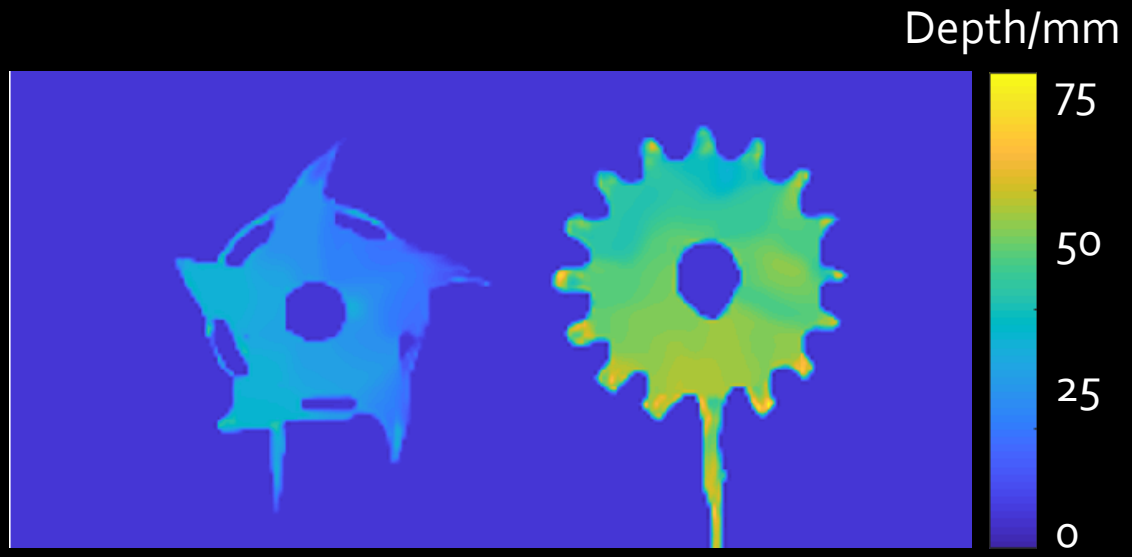
Applications: Intensity Imaging

Integration time 5.2ms

Applications: Depth Imaging



SPAD measurement
summed over time axis



Reconstruction depth map

Applications: Transient Imaging

Scene

SPAD measurement
(time resolved 20ps)

Reconstruction result

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
点昀 POINT SPREAD

Step Three: Enabling Comprimization for Color Channels

Learning Rank-1 Diffractive Optics for Single-shot High Dynamic Range Imaging

Limited Sensor Dynamic Range

Over Exposured

Under Exposured

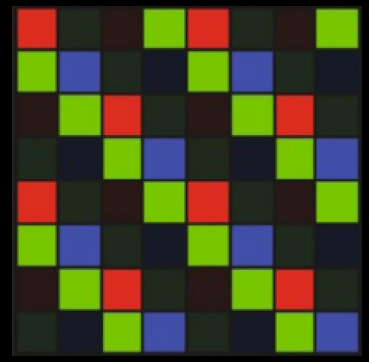
Traditonal Method—Burst Imaging

- Motion artifacts
- Long capture and processing time

Open Problem in HDR Imaging

Burst exposures

[Wan12]



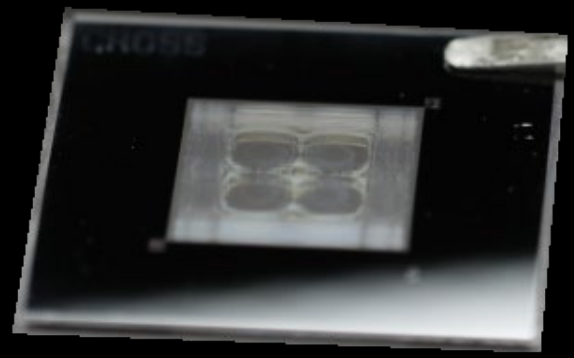
Deep L

Open Problem in HDR Imaging

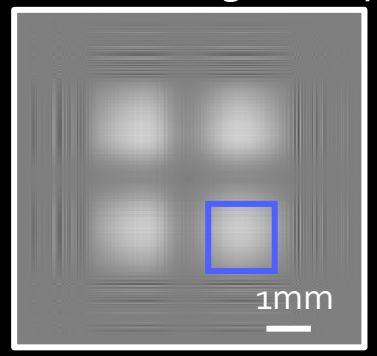
Single-shot HDR imaging using jointly learned optic + neural network

Prototype

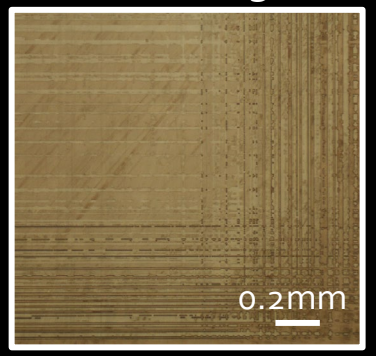
Fabricated learned optic



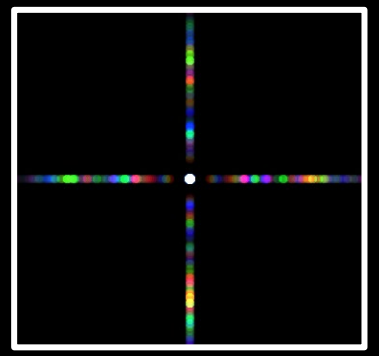
Simulated Height Map



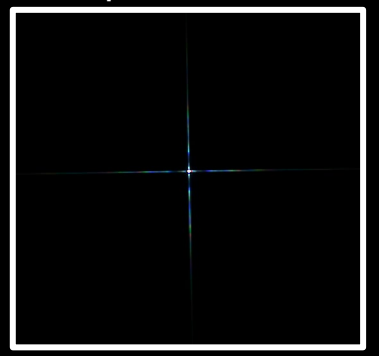
Fabricated Height Map



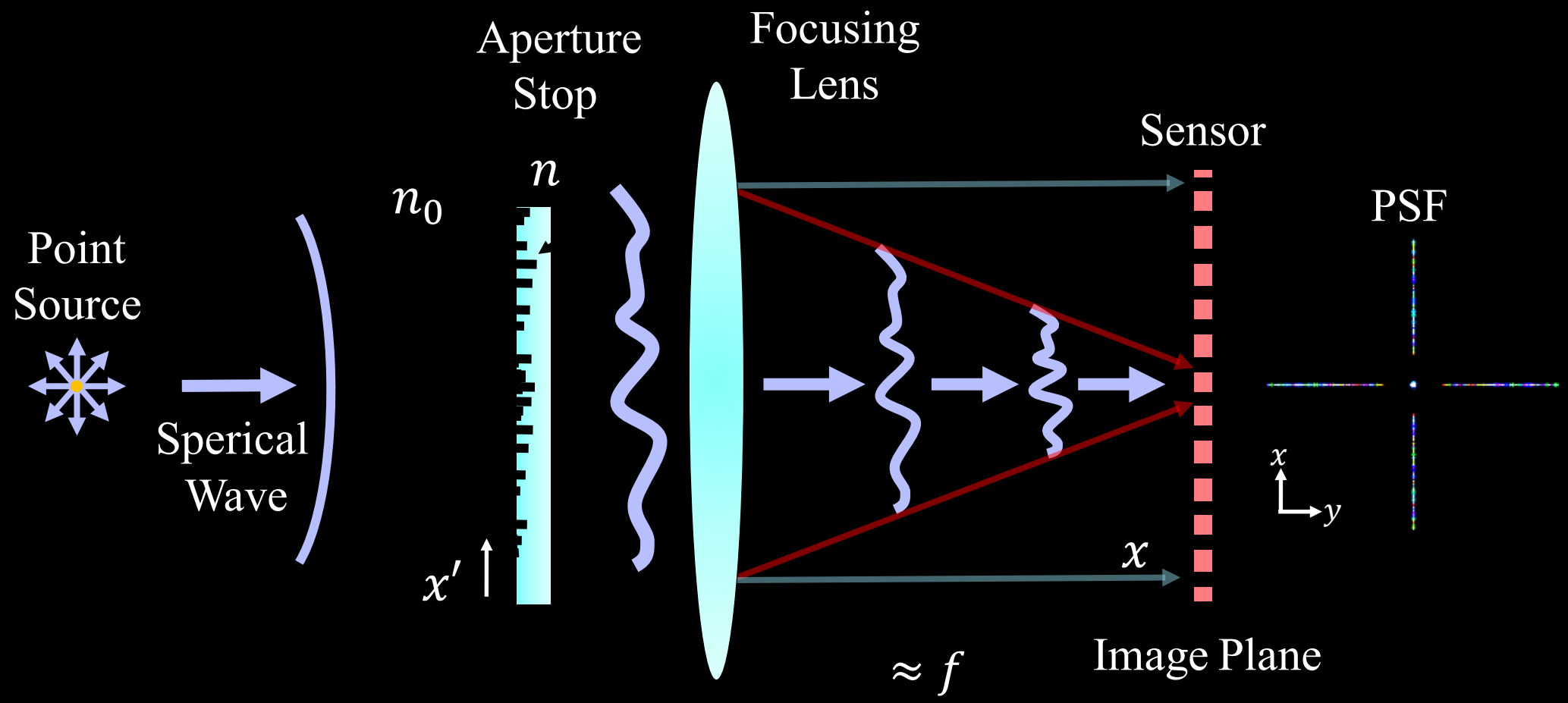
Simulated PSF



Captured PSF



Color-comprimised Optical Model



End-to-End HDRI Pipeline

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
点明 POINT SPREAD

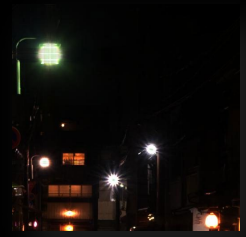


Image Formation Model

PSF

Sensor Model

$\mathbf{I}$

$\mathbf{I}_s$

Feature Extraction

VGG Features

Dilated Full-Resolution Reconstruction Block

Recovered Unsaturated Region

$\mathcal{L}_{\mathbb{U}}$

Residual

HDR Fusion

conv

conv

concat

conv

conv

conv

conv

$\mathcal{L}_{\mathbb{F}}$

Residual-Refinement

Recovered Saturated Region

$\mathcal{L}_{\mathbb{S}}$

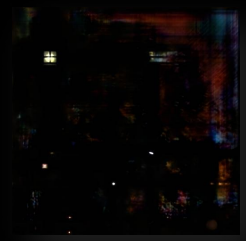
Phase

Capture

Prototype

Fabrication

Rank-1 Factorization



Learning Rank-1 Diffractive Optics

- Point Light Source

$$\mathbf{u}_{-} = A_0 e^{jk\sqrt{x'^2+y'^2+z^2}}$$

- DOE Layer and Rank-1 Factorization

$$\mathbf{u}_{+} = \mathcal{A}(x', y')\mathbf{u}_{-} e^{jk(n_\lambda - 1)\mathbf{h}(x', y')}$$

- Heightmap

$$\mathbf{h}(x', y') = h_{\max}\, \sigma\left(\mathbf{v}\mathbf{q}^{\top}\right)$$

Trainable parameters: $\mathbf{v} \in \mathbb{R}^{m\times 3}$ $\mathbf{q} \in \mathbb{R}^{m\times 3}$

- Fresnel Propagation

$$\mathbf{u}_s = \mathcal{F}^{-1}\{\mathcal{F}\{\mathbf{u}_l\}\mathcal{H}\}$$

$$\mathcal{H}(f_x, f_y) = e^{jkL} e^{-j\pi\lambda L(f_x^2+f_y^2)}$$

Loss Function

- Total Loss

$$\mathcal{L}_{\text{Total}} = \mathcal{L}_{\mathbb{U}} + \mathcal{L}_{\mathbb{S}} + \mathcal{L}_{\mathbb{F}}$$

- Loss on Unsaturated Regions
  - VGG Loss $\mathcal{L}_{\text{VGG}}(\hat{x}, x) = \sum \nu_l \|\phi_l(\hat{x}) - \phi_l(x)\|_1$
  - Exclusion loss $\mathcal{L}_{\text{excl}} = \|\tanh(\eta_{\mathbb{U}} \mid \nabla \hat{\mathbf{I}}_{\mathbb{U}} \mid) \odot \tanh(\eta_r \mid \nabla \hat{\mathbf{I}}_r \mid)\|_F$
- Loss on Saturated Region

$$\mathcal{L}_{\mathbb{S}} = \beta \mathcal{L}_{\text{VGG}}(\hat{\mathbf{I}}_{\mathbb{S}}, \mathbf{I}_{\mathbb{S}})$$

- Loss in Fused Output

$$\mathcal{L}_{\mathbb{F}} = \mathcal{L}_{\text{Huber}}\left((\hat{\mathbf{I}} + \epsilon)^{\gamma}, (\mathbf{I} + \epsilon)^{\gamma}\right)$$

Simulated Results

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

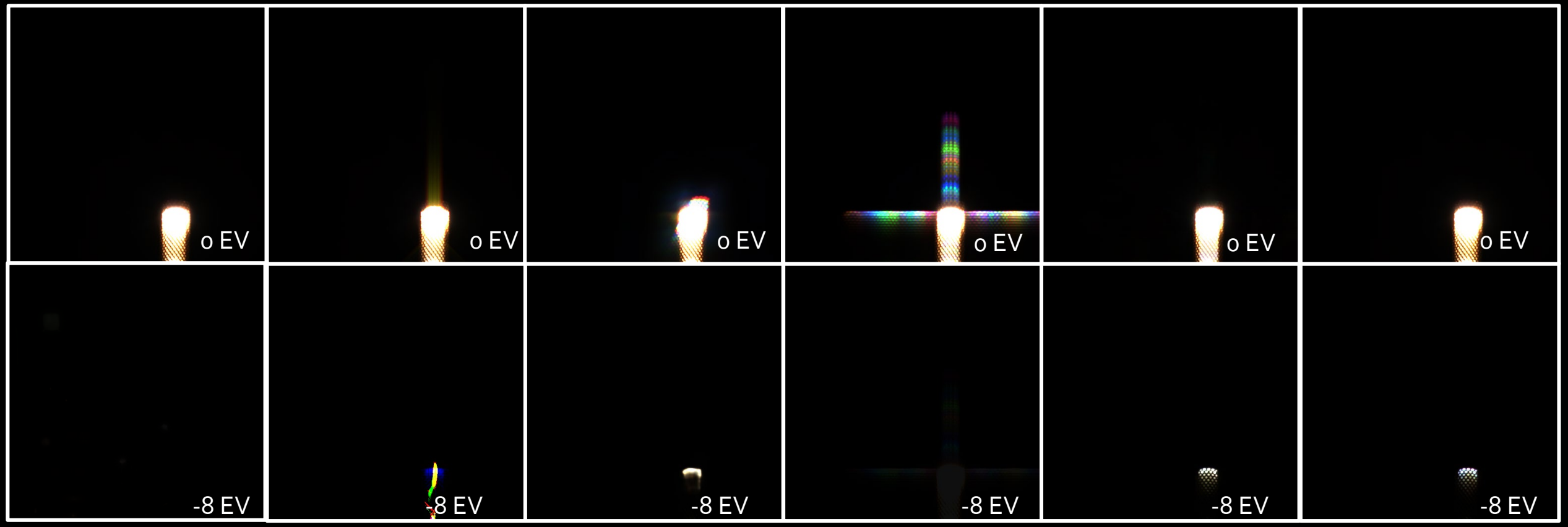
点明 POINT SPREAD

HDR-CNN | Glare-HDR | Deep Optics | Our Measurement | Our Reconstruction | Reference HDR

Simulated Results

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

点明 POINT SPREAD



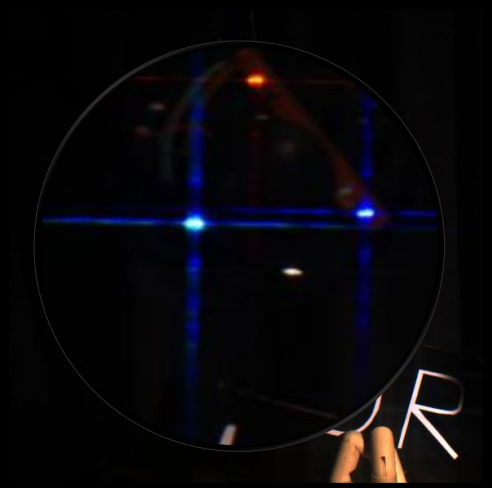
Comparison in Simulation

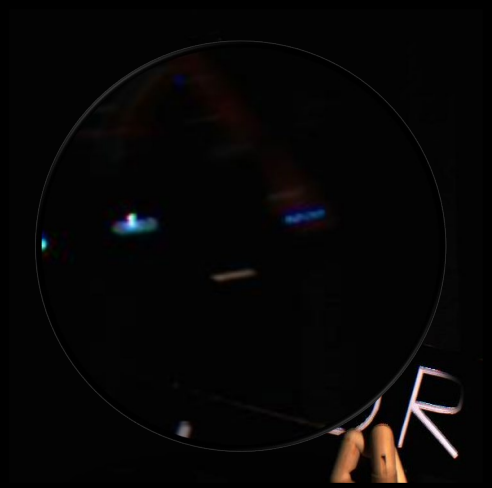
香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen

点昀 POINT SPREAD

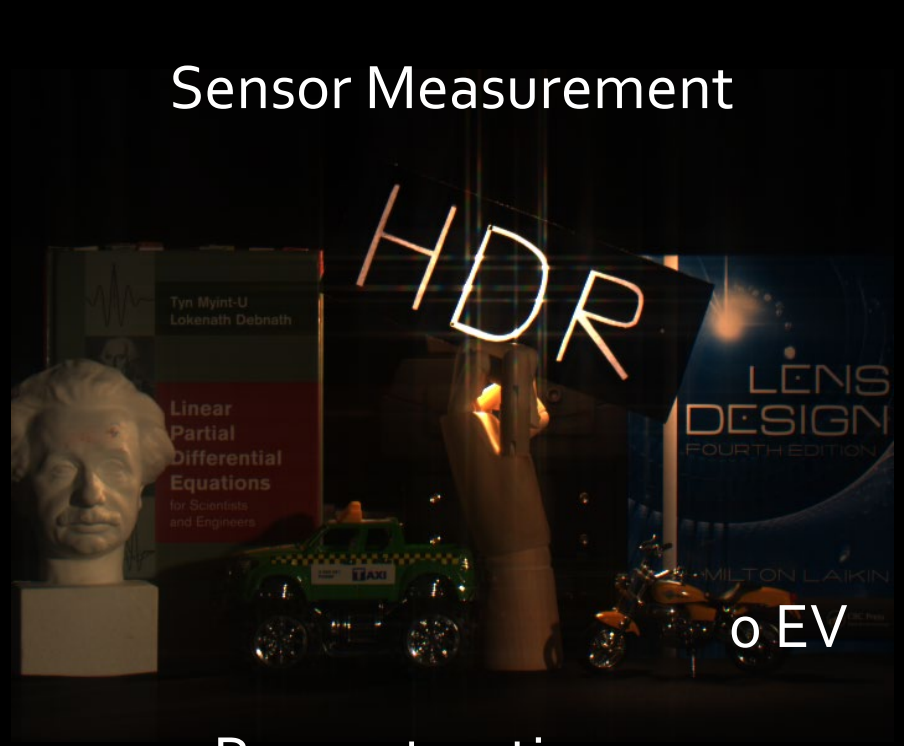
Methods	PSNR	HDR-VDP 2
Ours	48.26	74.47
Deep Optics	40.30	67.96
Glare-HDR	32.23	56.76
HDR-CNN	34.06	54.34
LDR	33.57	52.43





Sensor Measurement



Reconstruction

Measurement	Reconstruction	Reference
-4 EV	-4 EV	-4 EV
-6 EV	-6 EV	-6 EV

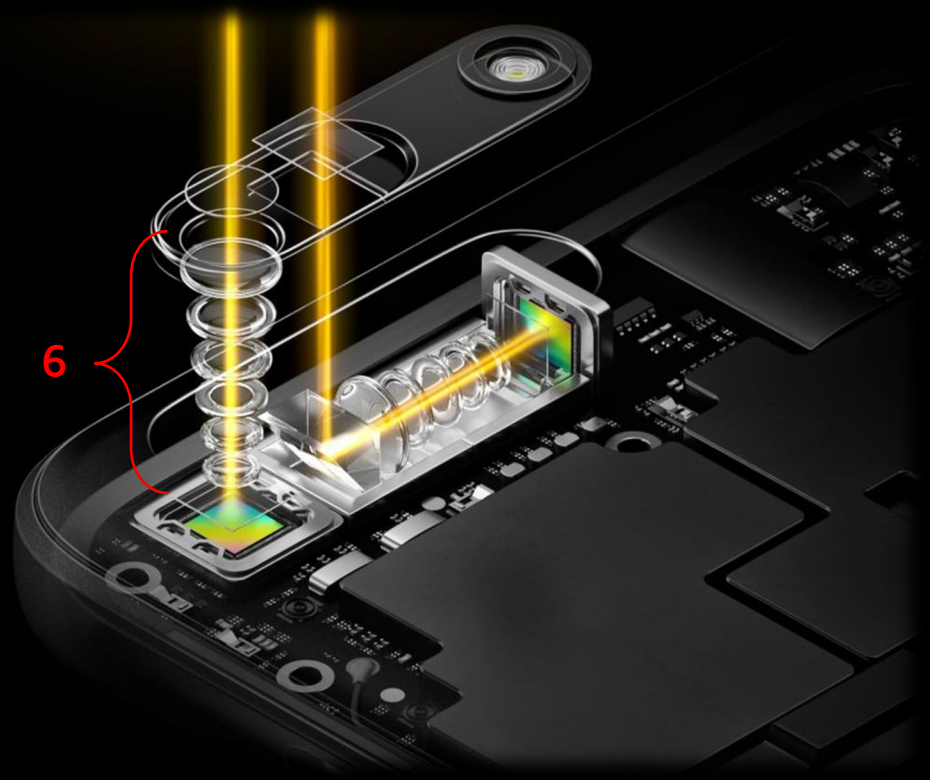
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
点昀 POINT SPREAD

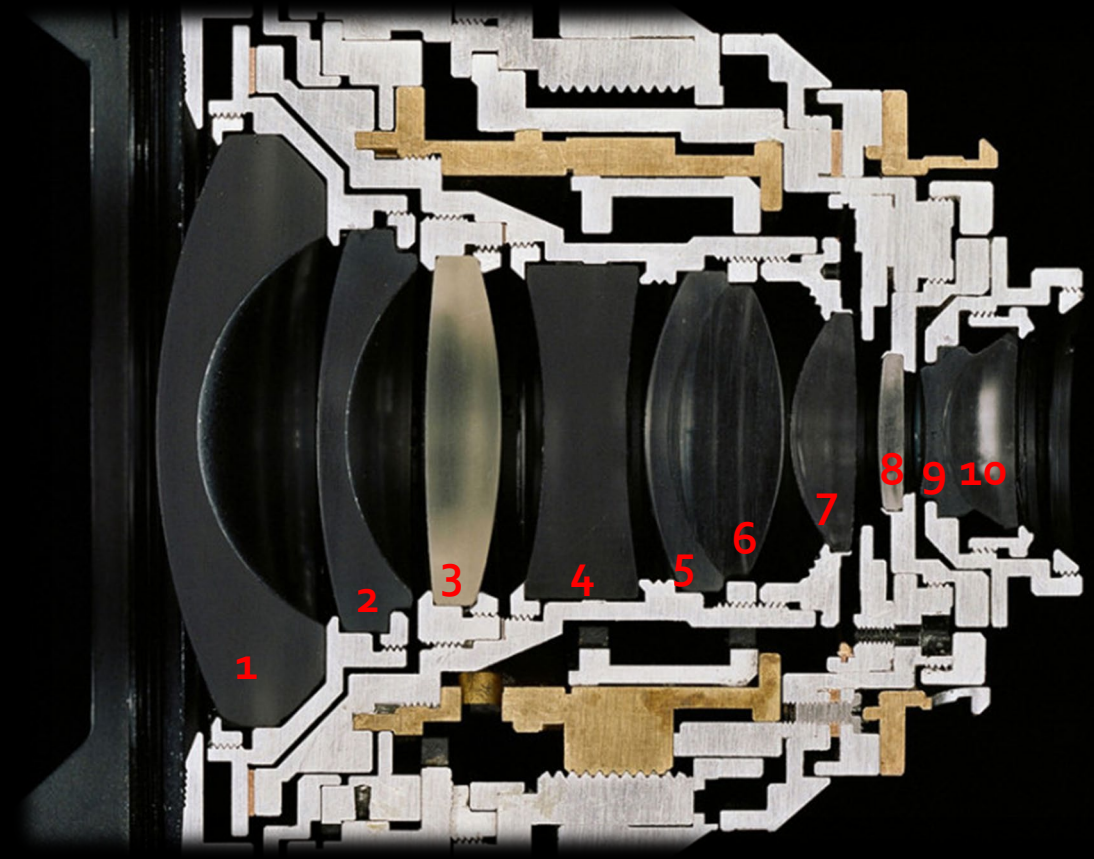
Measured	Recovered	Reference	Measured	Recovered	Reference
					
					
					

Step Four: End-to-end Complex Lens Design

End-to-End Complex Lens Design with Differentiable Ray Tracing

Complex Lens Systems





- High quality imaging combat with aberrations of all kinds
- Sometimes not optimal for a given task

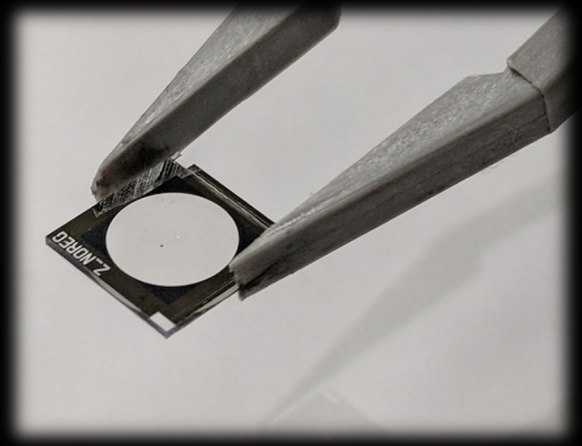
** Images credits @ Oppo, Nikon*

End-to-end Optics Designing

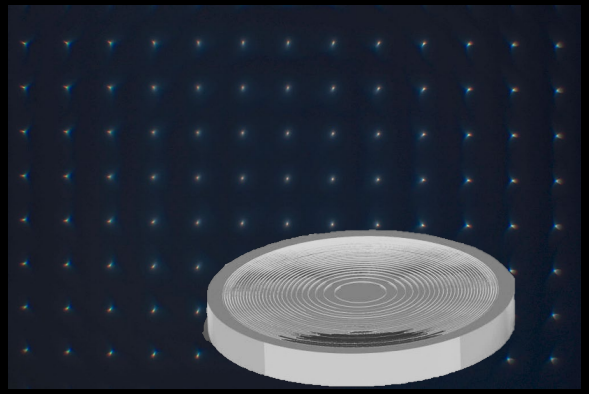
- Diffractive optics
- Refractive optics

- Thin form factor

- Single layer differentiable
- Paraxial approximation
- Limited design freedom
- Limited image quality
- Not fully differentiable

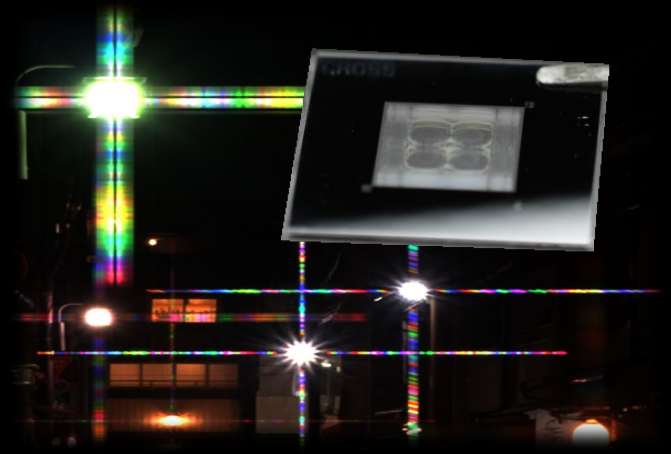


[Sitzmann et al. 2018]



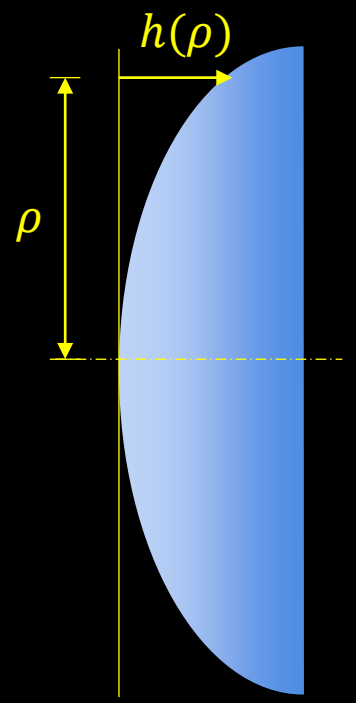
[Peng et al. 2019]

[Sun et al. 2020]



[Sun et al. 2020]

Differentiable Lens Model and Ray Tracer



- Aspherical Lenses:

$$h(\rho) = \frac{c\rho}{1+\sqrt{1-\alpha\rho}} + \sum_{i=2}^{n} a_{2i}\rho^{i} \qquad \alpha = (1+\kappa)c^2$$

- Aspherical Lenses's Derivative:

$$h'(\rho) = \mathrm{c}\frac{1+\sqrt{1-\alpha\rho}-\alpha\rho/2}{\sqrt{1-\alpha\rho}(1+\sqrt{1-\alpha\rho})^2} + \sum_{i=2}^{n} i a_{2i}\rho^{i-1}$$

- Implicit form $f(x,y,z)$ and its derivative:

$$f(x,y,z) = h(\rho) - z$$

$$\nabla f = (2h'(\rho)x, 2h'(\rho)y, -1)$$

Differentiable Lens Model and Ray Tracer

- Ray-surface Intersection by Newton's Method: Ray(**o, d**)
  - Find the t>o such that:

$$f(x, y, z) = f(\mathbf{o} + t\mathbf{d}) = 0$$

  - We solve the problem numerically using Newton's method

$$t^{(k+1)} \leftarrow t^{(k)} - \frac{f(\mathbf{o} + t^{(k)}\mathbf{d})}{f'(\mathbf{o} + t^{(k)}\mathbf{d})}$$

$$\leftarrow t^{(k)} - \frac{f(\mathbf{o} + t^{(k)}\mathbf{d})}{\nabla f \cdot \mathbf{d}}$$

- Dispersion extention of Mitsuba2 by Cauchy's equation

$$n(\lambda) = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \ldots \qquad A = n_D - \frac{B}{\lambda_D^2} \qquad B = \frac{n_D - 1}{V(\lambda_F^{-2} - \lambda_C^{-2})} \quad \ldots$$

Differentiable Lens Model and Ray Tracer

- Recorded image on sensor:

$$I_c(x', y') = \int Q_c(\lambda)[p(x', y', d, \lambda) * s_c(x', y', d)]d\lambda + n(x', y')$$

- Monto Carlo Sampling
- Image auto alignment during training:
  - The relationship between traced points on sensor $\boldsymbol{r}_d$ and ideal points $\boldsymbol{r}$ can be expressed as:

$$\boldsymbol{r}_d = \zeta \boldsymbol{r}(1 + k_1\boldsymbol{r}^2 + k_2 4 + k_3\boldsymbol{r}^6)$$

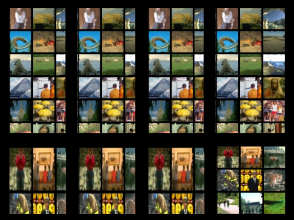
  - We solve a least square problem to obtain the current distortion coefficients and magnification.

$$\min_{\boldsymbol{K}} \left\| [\boldsymbol{r}, \boldsymbol{r}^3, \boldsymbol{r}^5, \boldsymbol{r}^7]\boldsymbol{K}^T - \boldsymbol{r}_d \right\|_2^2$$

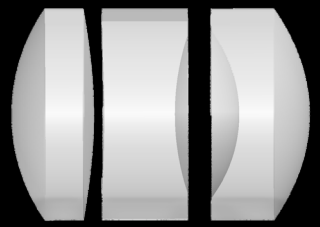
Differentiable Complex Lens and Recovery

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
点明 POINT SPREAD

Scene Dataset

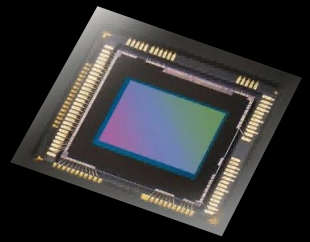


Complex Lens Model

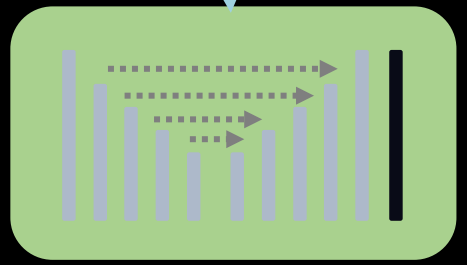


differentiable ray-tracing

Simulated Sensor



Loss



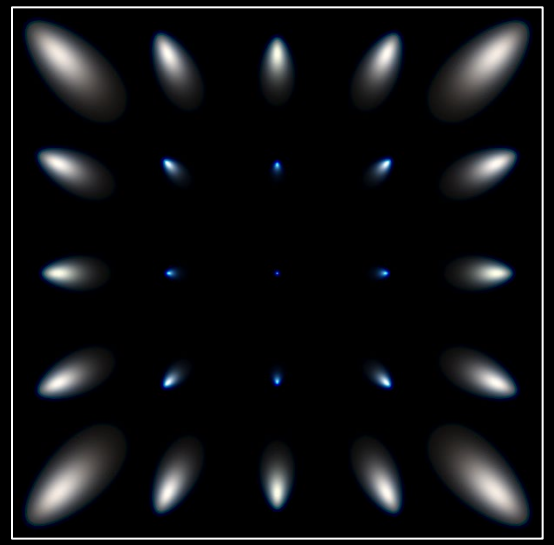
Real Scene

Prototype

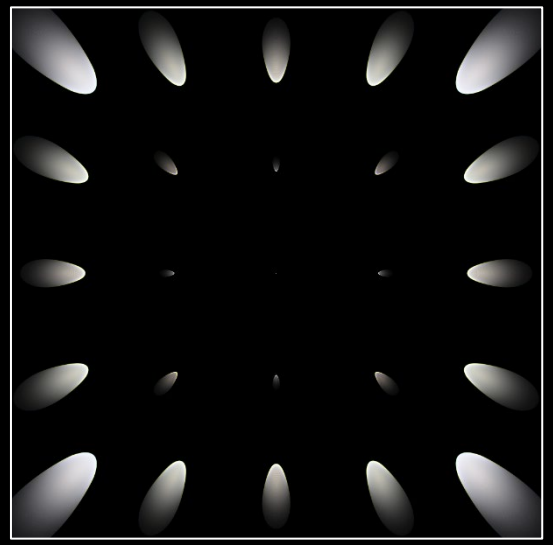
Image Recovery Network

Large Field-of-View Imaging

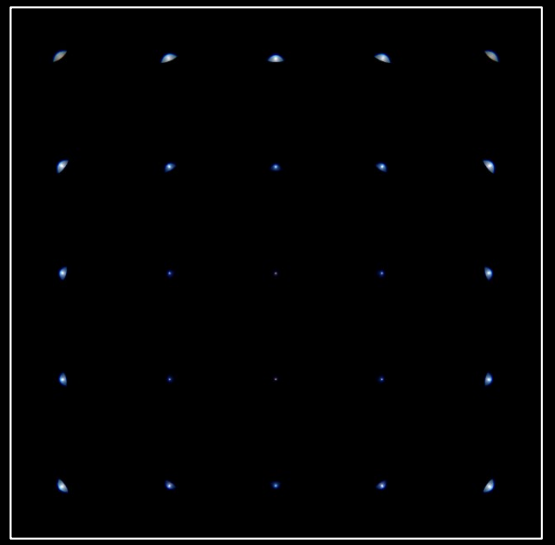
AL2550



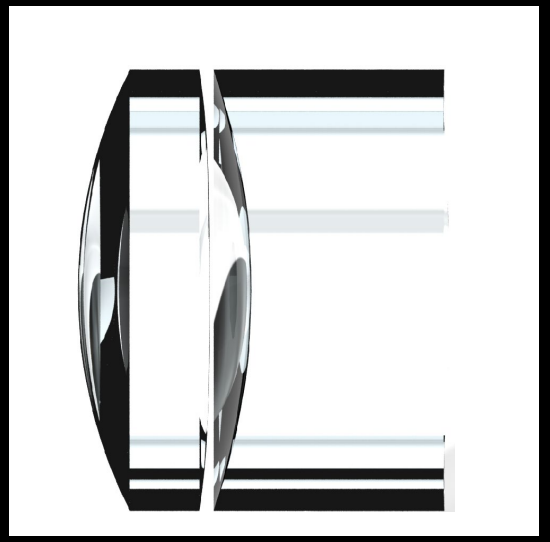
ACA254

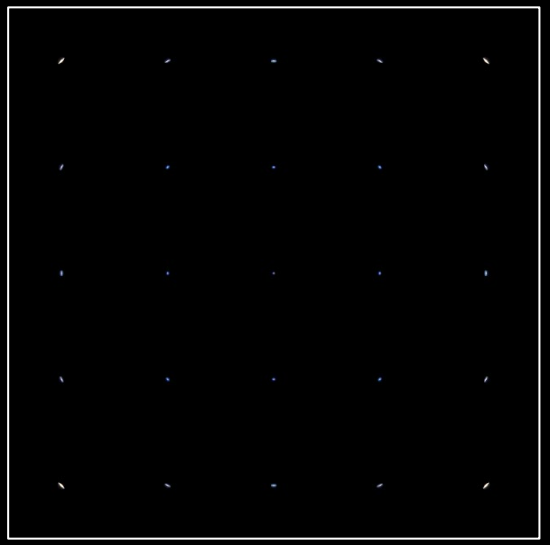


Peng et al 2019



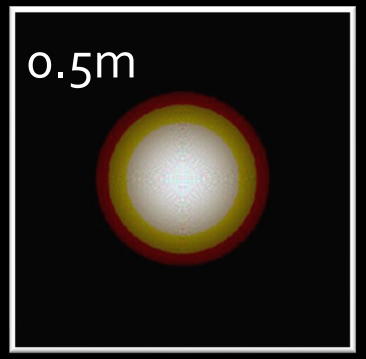
Ours

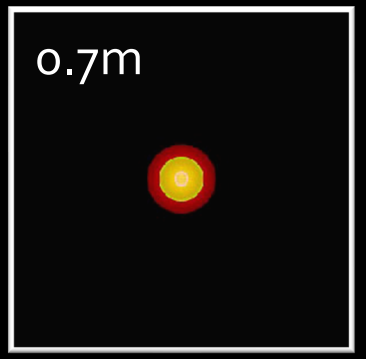


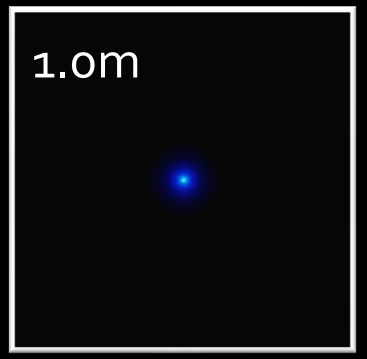


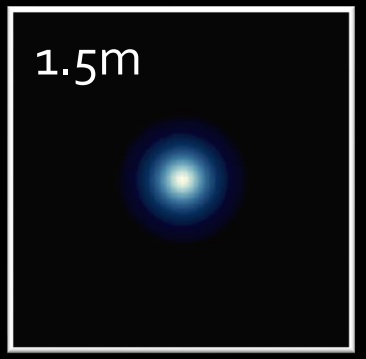
Extended Depth-of-Field

AL2550

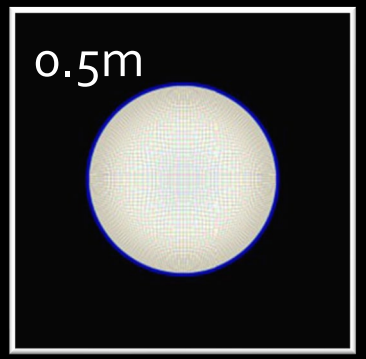


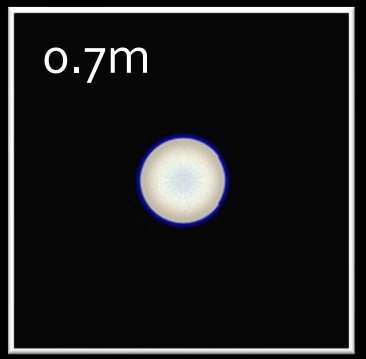


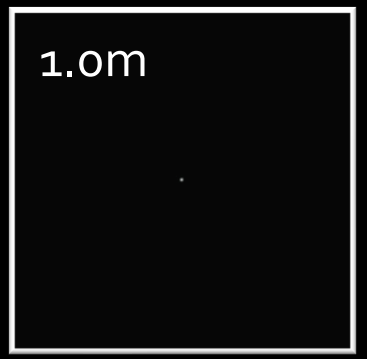


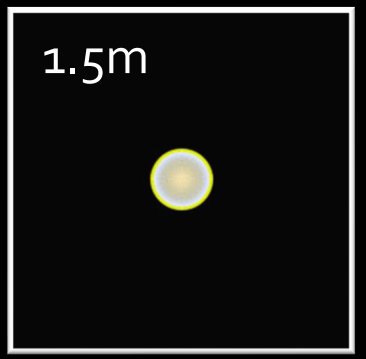


ACA254

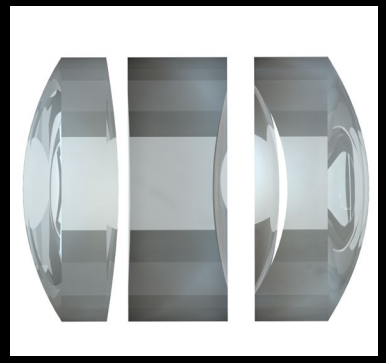


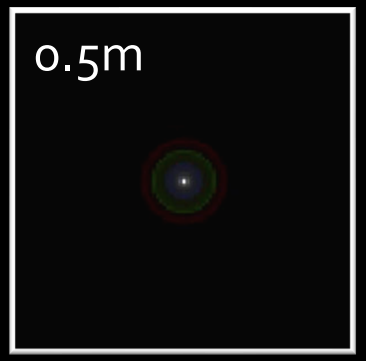


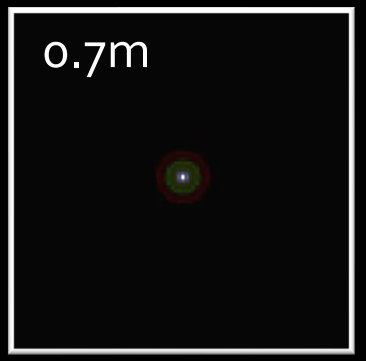




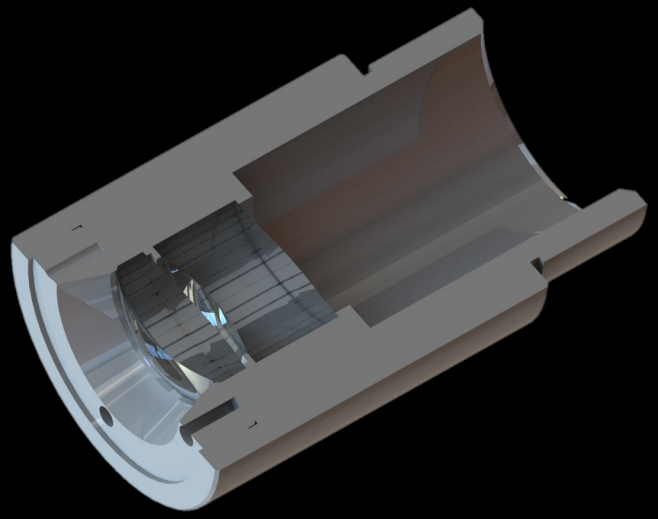
Ours



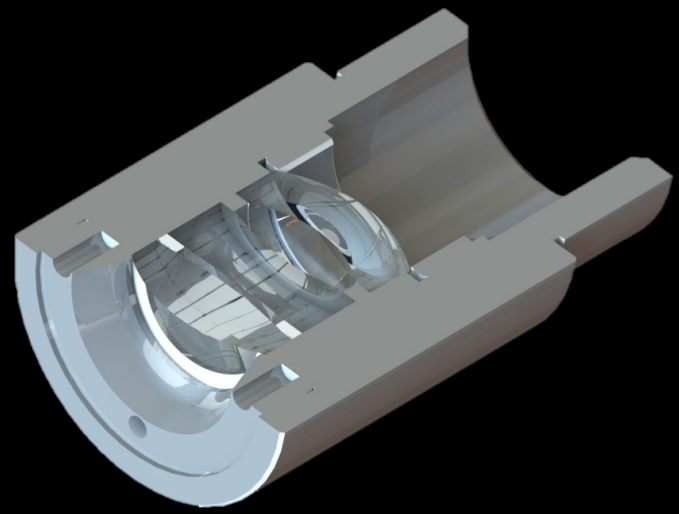




Prototypes

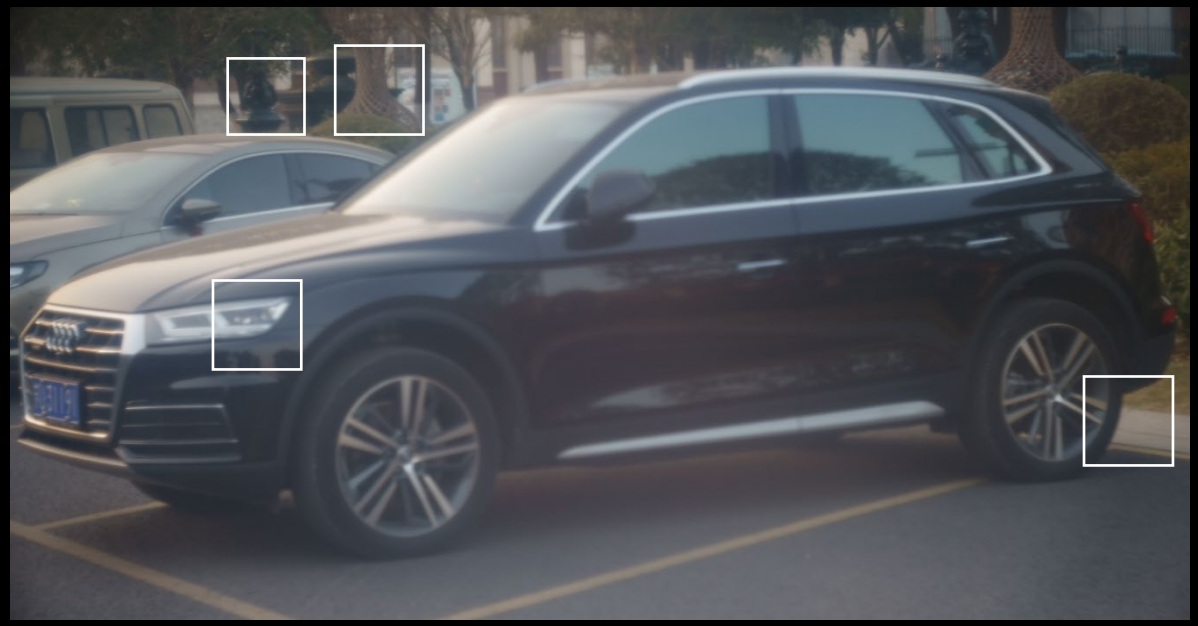


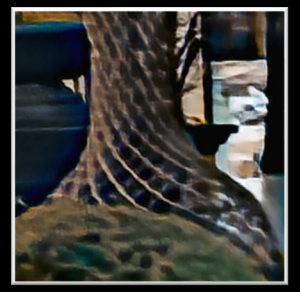
50mm/f4



50mm/f4

Results of Large Field-of-View Imaging

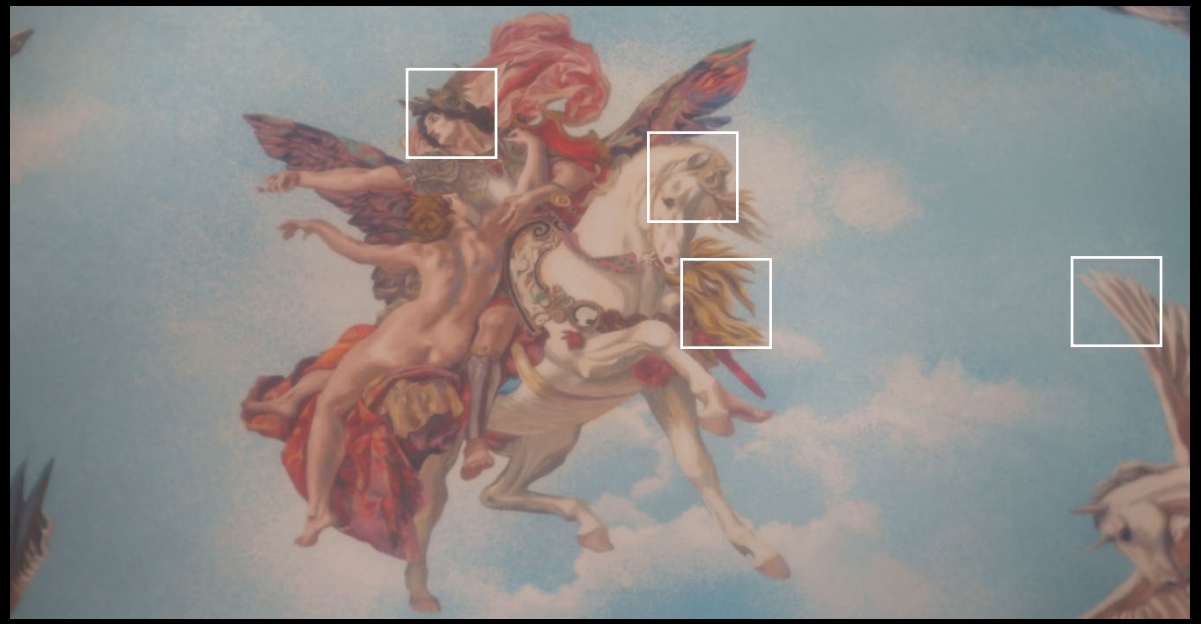


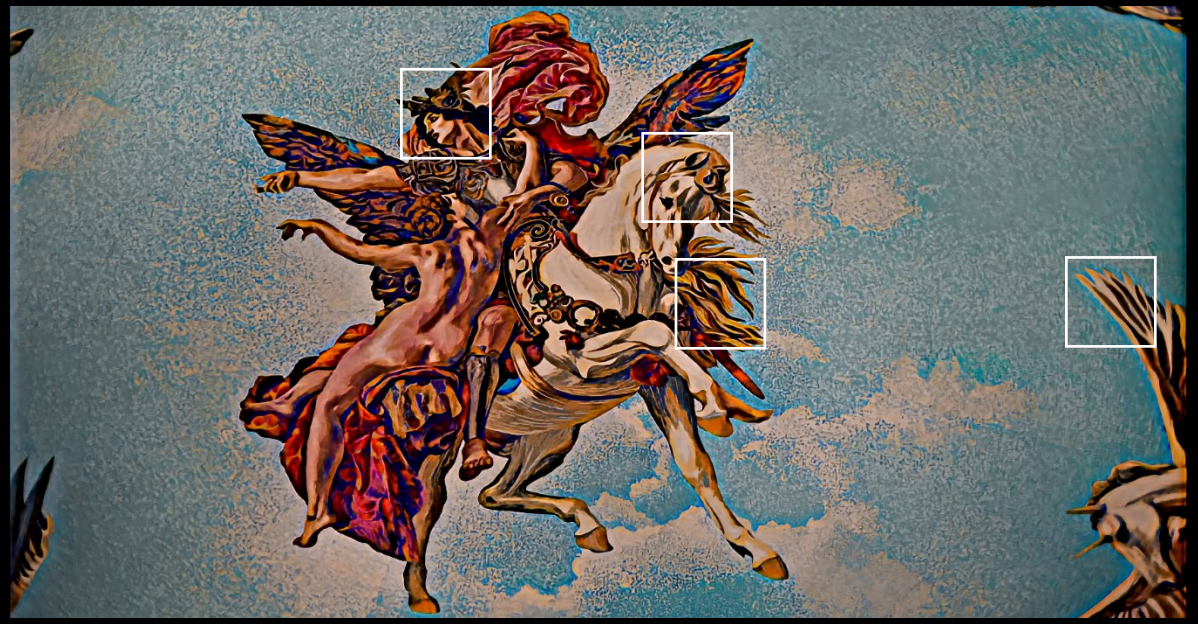


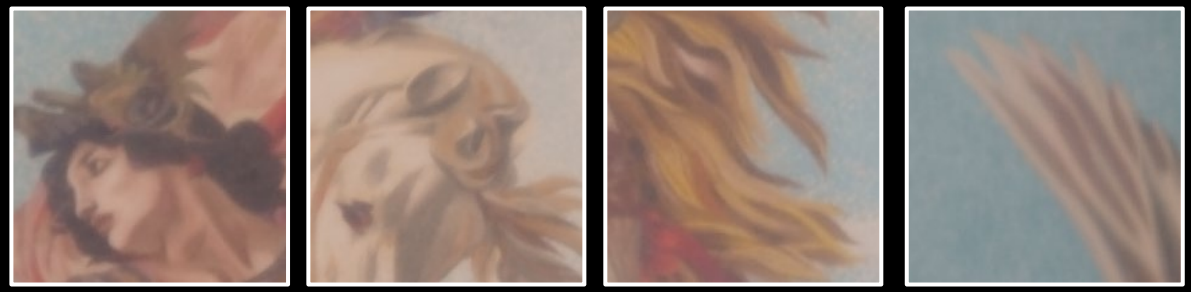
Captured

Recovered

Results of Large Field-of-View Imaging





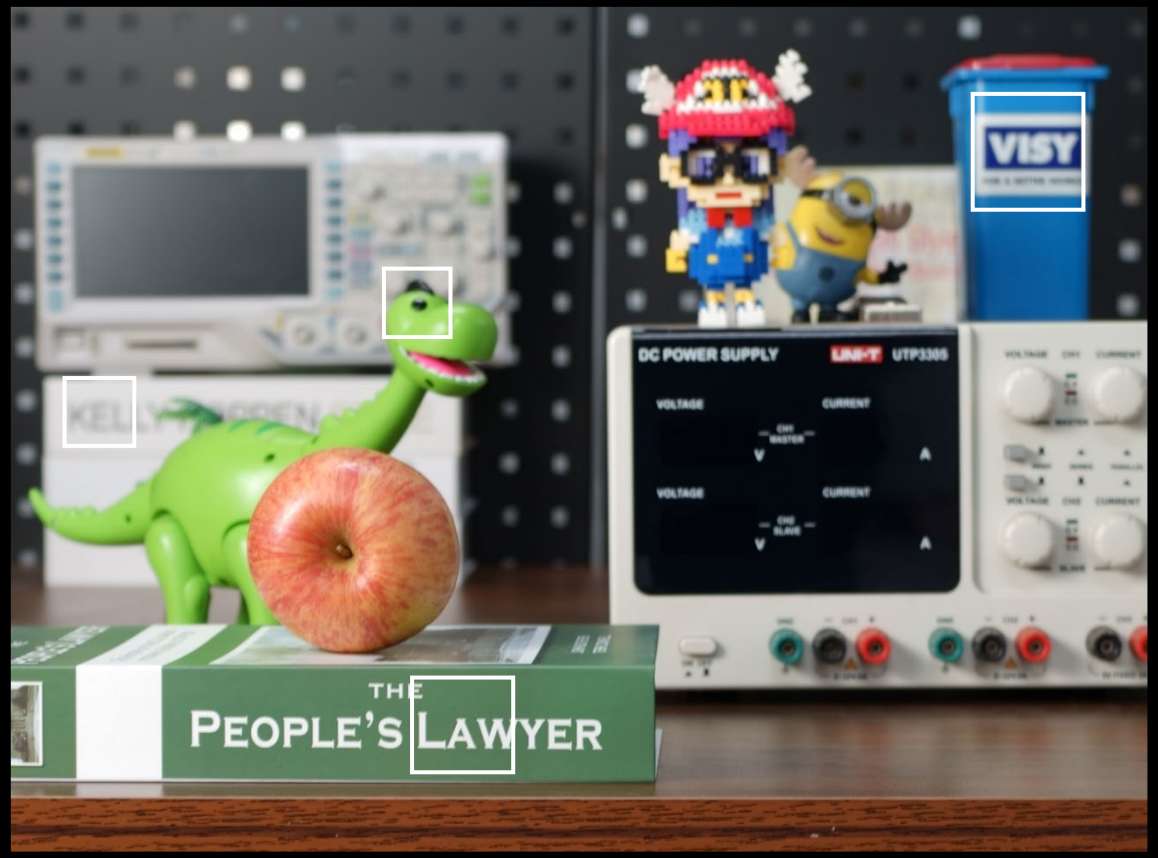


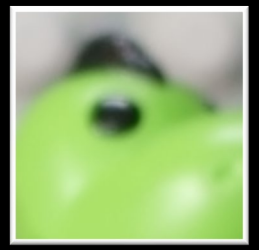
Captured

Recovered

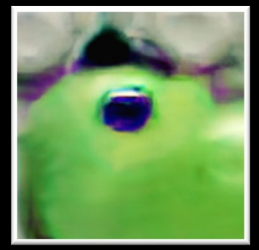
Results of Extended Depth-of-Field

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
点明 POINT SPREAD





Sony standard zoom lens at *50mm/f4.5*



Ours at *50mm/f4*

Results of Extended Depth-of-Field

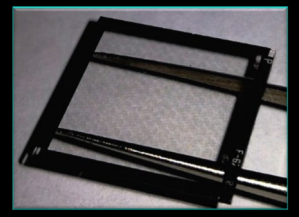
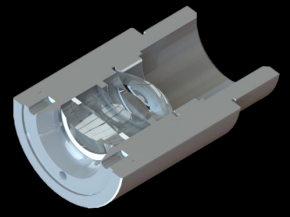
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen
点明 POINT SPREAD

Ours at *50mm/f4*
Captured

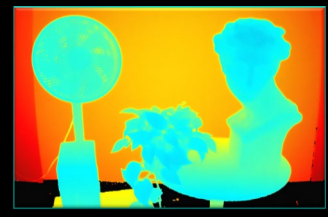
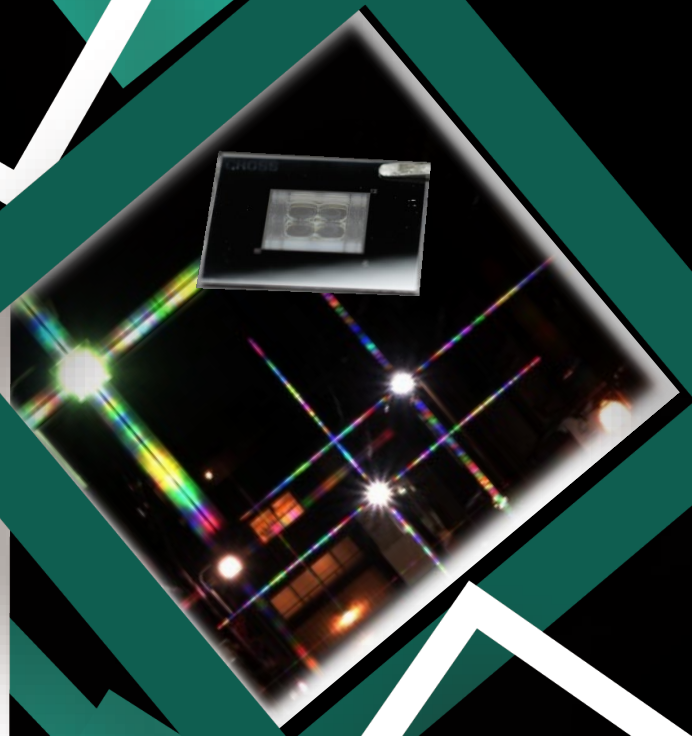
Ours at *50mm/f4*
Processed

Sony standard zoom lens
at *50mm/f4.5*

- Step One: Joint Optimization of Optics and Image Recovery
  - Learned Large Field-of-View Imaging with Thin-Plate Optics
- Step Two: Compressive and Optically Coded Super-resolution SPAD Camera
  - Compressive and Optically Coded Super-resolution SPAD Camera
- Step Three: Enabling Comprimization for Color Channels
  - Learning Rank-1 Diffractive Optics for Single-shot High Dynamic Range Imaging
- Step Four: End-to-end Complex Lens Design
  - End-to-end Complex Lens Design with Differentiable Ray-Tracing



Computational Imaging

Thank You!

Qilin Sun（孙启霖）

香港中文大学（深圳）

点昀技术（Point Spread Technology）