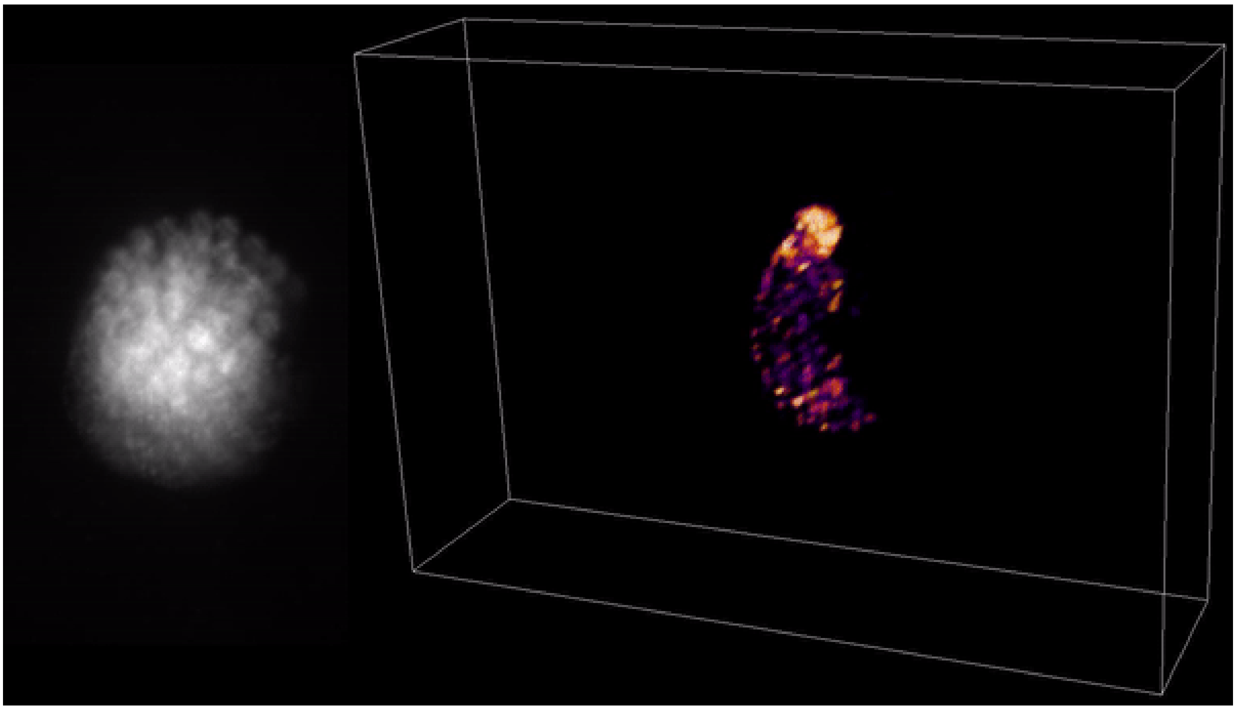
EE 046746 - Technion - Computer Vision

Elias Nehme

Tutorial 14 - Deep Computational Imaging



* Image source

Agenda

What is Computational Imaging?

Computational Imaging

Co-design of optics, sensing, and algorithms!

Computational Imaging

What is Computational Imaging?

Computational Imaging

HDR Imaging [Mann,Devebec,Nayar,…]

Super-resolution [Baker,Ben-Ezra,…]

EDOF [Dowski,Nayar,…]

Light Fields [Levoy,…]

Compressive Imaging [Baraniuk,…]

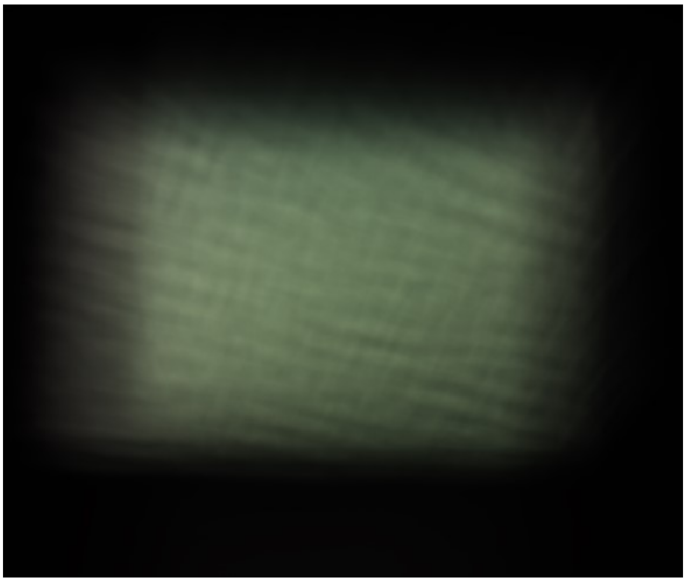
Computational Imaging

Compressive Imaging

- Depth encoding PSF

Depth Encoding PSF

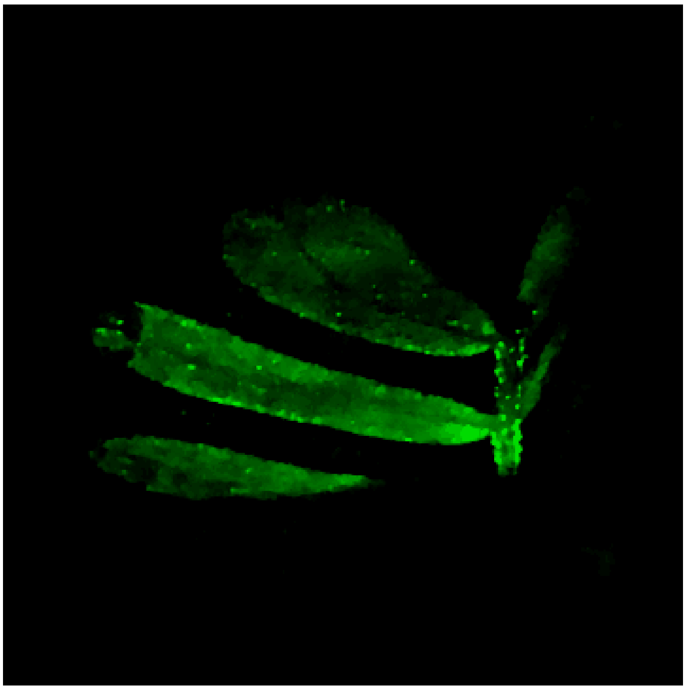
- Measurement is a 2D image:



- Image source - Optica 2018

Depth Encoding PSF

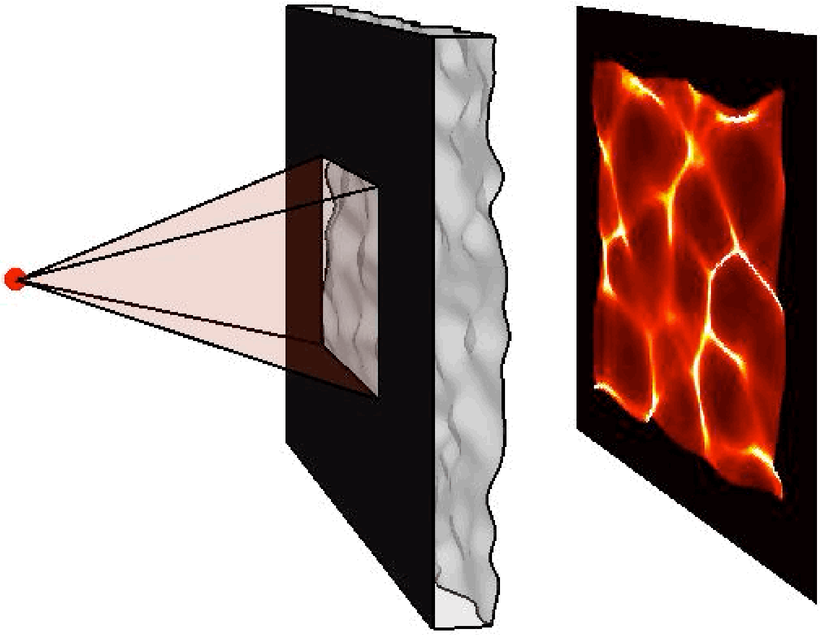
- Recovery is a 3D volume:



- What?? How?!
- Image source - Optica 2018

Depth Encoding PSF

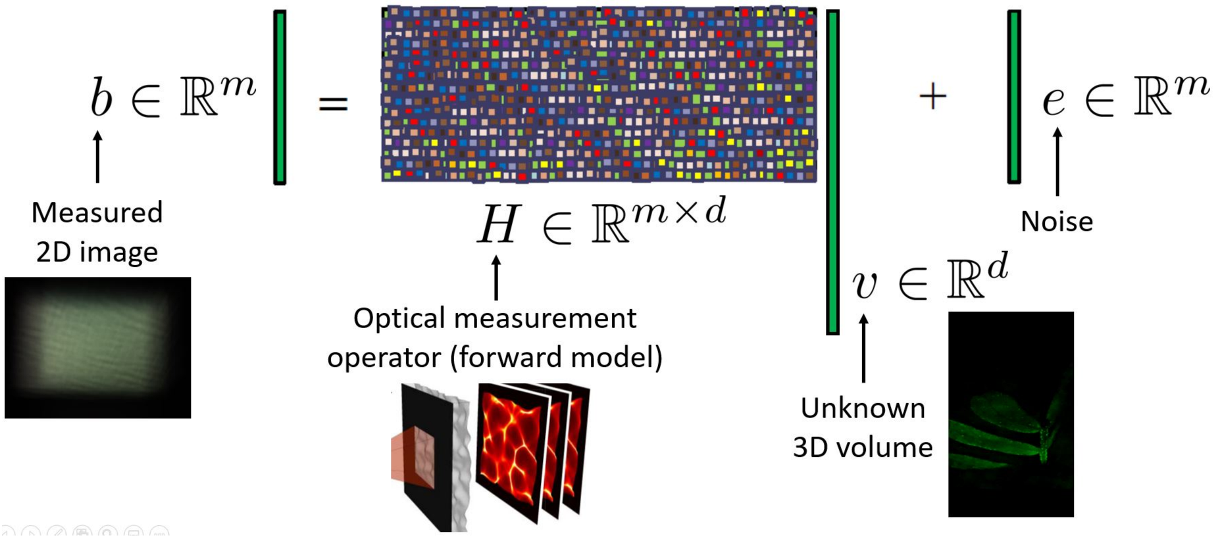
- Depth Encoding Impulse Response/Point Spread Function (PSF)
  - Main idea is to encode depth in the shape generated on the 2D sensor



- Image source - Optica 2018

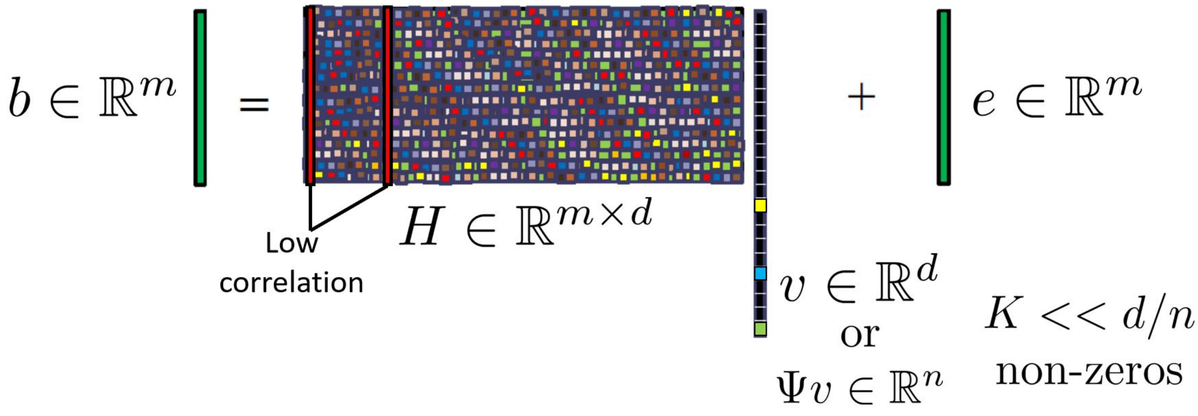
Depth Encoding PSF

- Writing down the problem in matrix formulation



Depth Encoding PSF

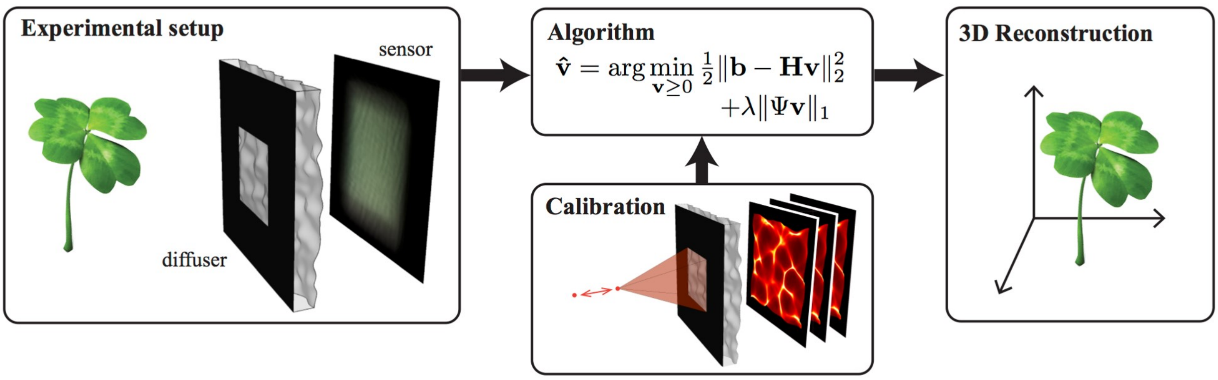
- Solution given by "MAP" estimator under certain conditions



$$\hat{v}_{TV} = argmin_{v\geq 0}\left\{\frac{1}{2}\|b - Hv\|^2 + \lambda\|\Psi v\|_1\right\}$$

Depth Encoding PSF

- Overall framework:



- Image source - Optica 2018

Deep Optics

- Computer Vision Pipelines
- Differentiable Optics

Computer Vision Pipelines

Computer Vision Pipelines

How do CV pipelines work?

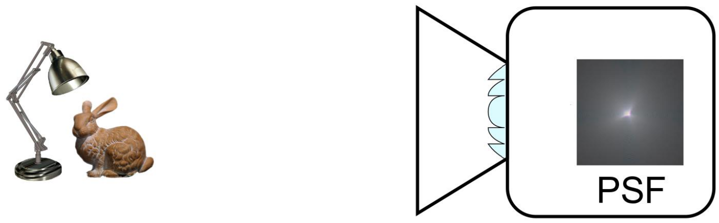
Computer Vision Pipelines

Objective: Solve CV problem on scene

What is this?

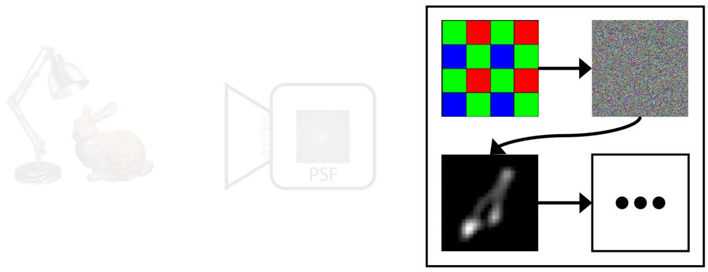
Computer Vision Pipelines

Step 1: Build camera



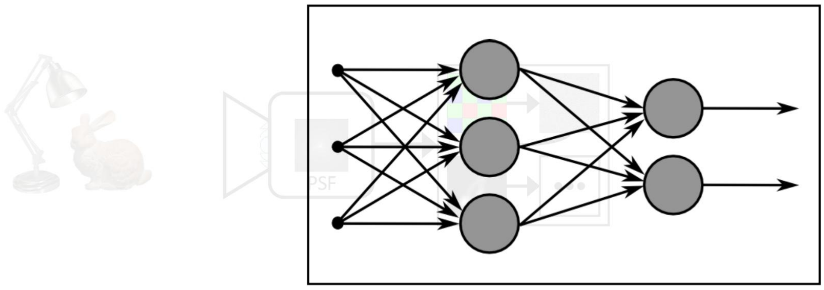
Optimize optics to minimize aberrations:
Blur/spot size, chromatic aberrations, distortions, …

Step 2: Image Signal Processing (ISP)



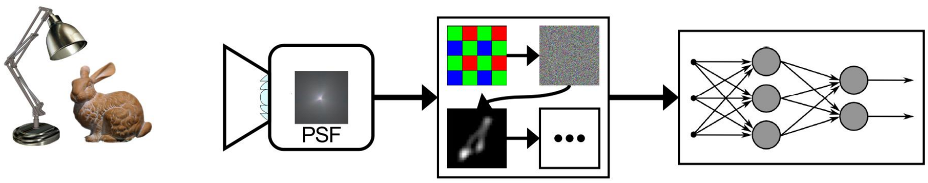
Maximize PSNR:
Demosaicking, Denoising, Deblurring, …

Step 3: CNN for Semantic task

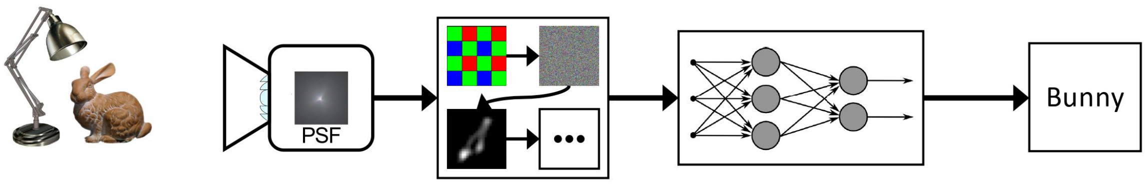


Minimize semantic loss:
classification error, segmentation error, …

…

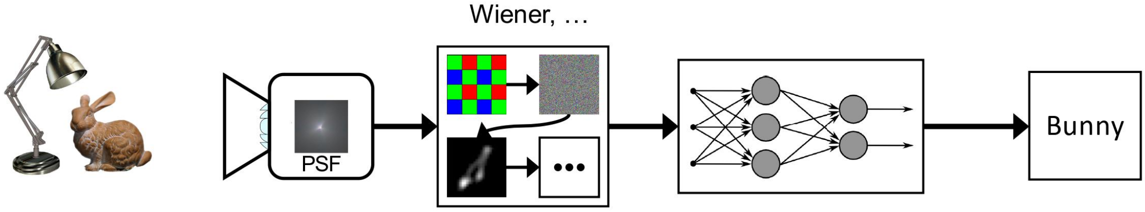


It works!



Prior work on optimizing each part of pipeline

BM3D, NLM, HQS,
Wiener, …



Wavefront coding,
Diffractive Achromat,
…

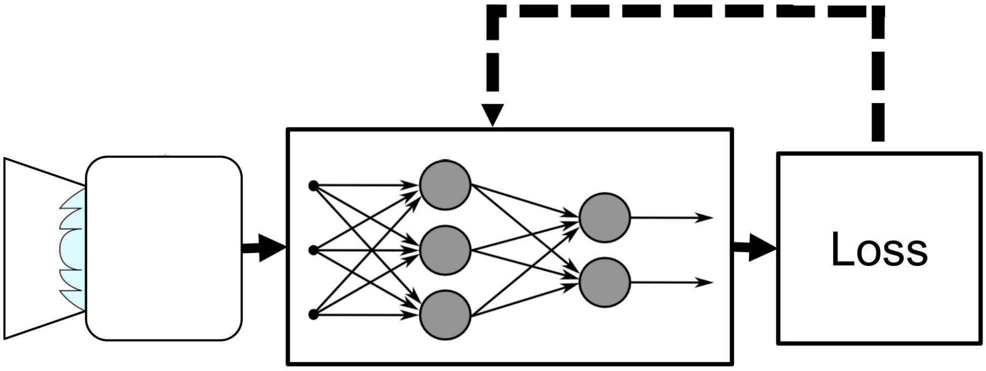
Alexnet, VGG19,
Yolo, Densenet,
…

Computer Vision Pipelines

- Animal vision is adapted to the surrounding and the day-to-day "task"

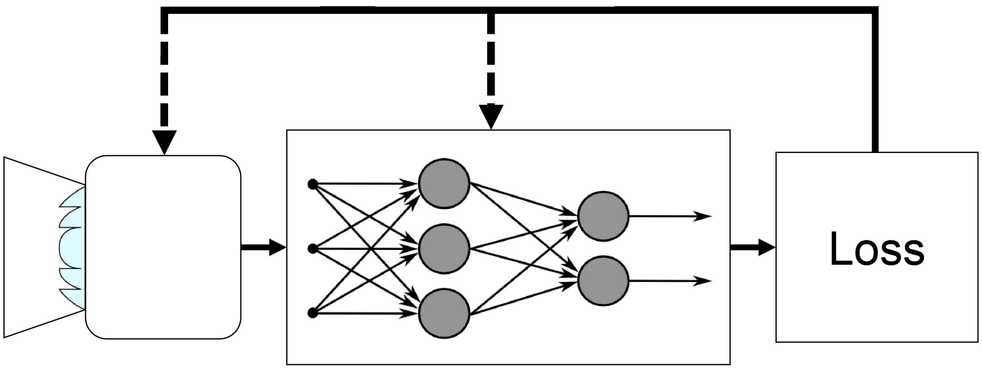
Computer Vision Pipelines

- "Standard" deep image processing



Computer Vision Pipelines

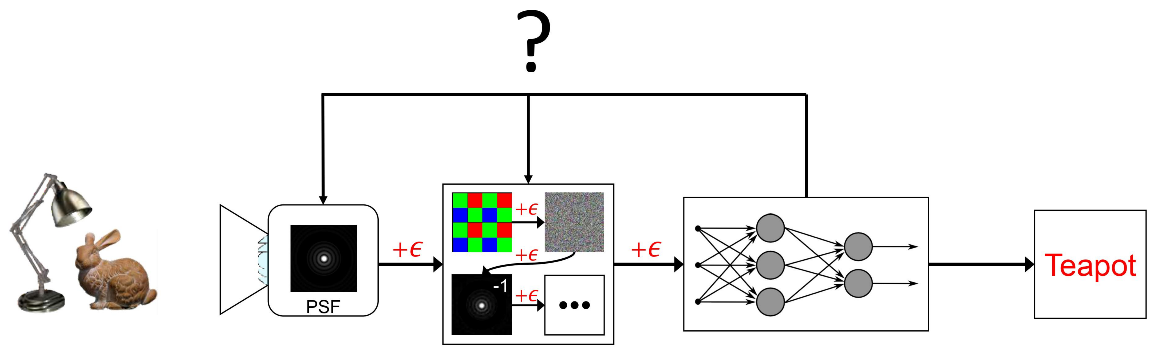
- Deep computational imaging



Optimize optics end-to-end with higher-level processing!

Computer Vision Pipelines

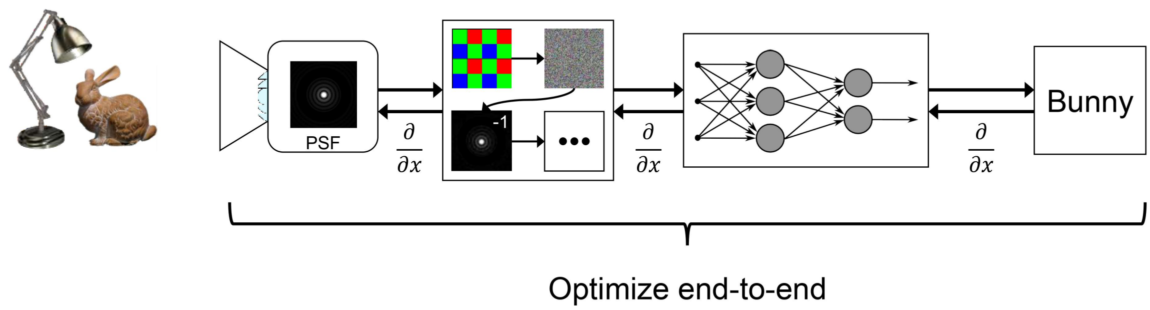
- Image classification with specialized "optics"



Computer Vision Pipelines

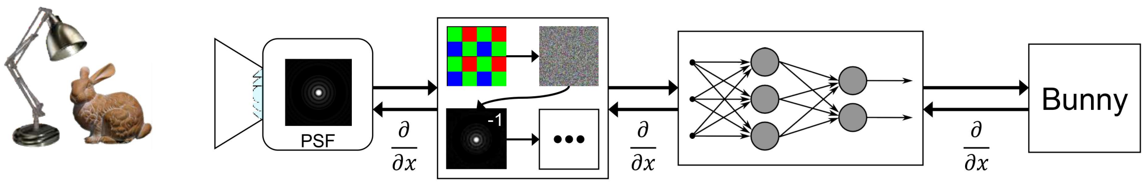
- Main idea: optimize the optics and the algorithm jointly to excel in the final task

Vision: The Deep Computational Camera

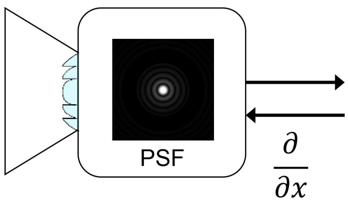


Computer Vision Pipelines

Vision: The Deep Computational Camera

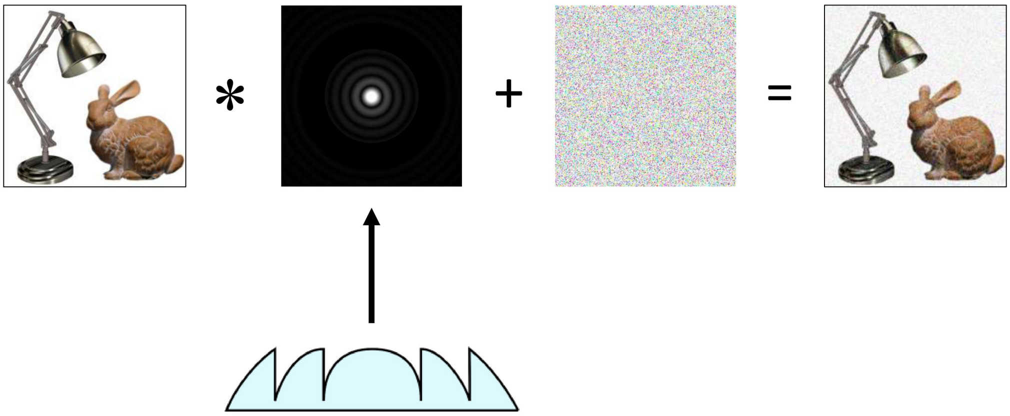


- Performance & robustness gains
- Domain-specific hardware may reduce footprint, cost, power…
- New design space: The "BunnyCam"



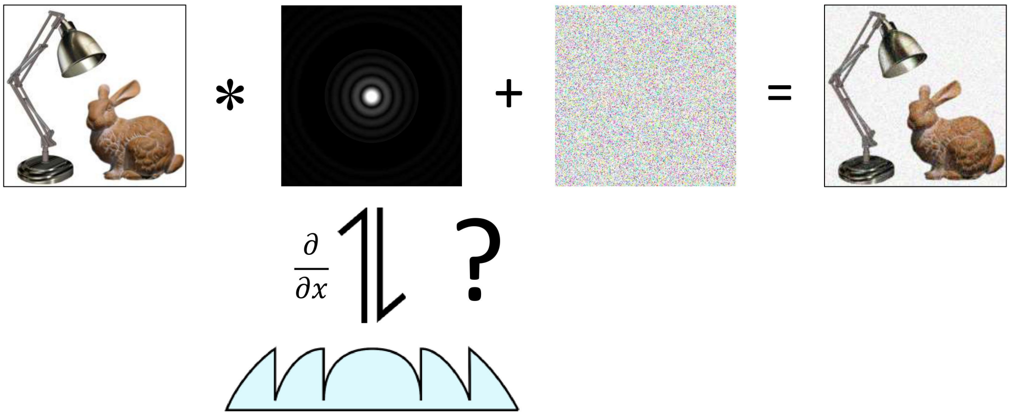
A differentiable optics model

Image formation model

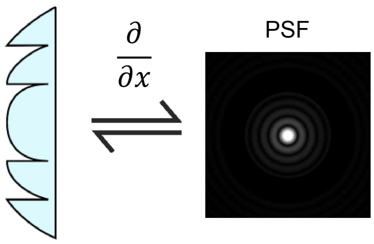


Differentiable Optics

How does the optical element map to the PSF?

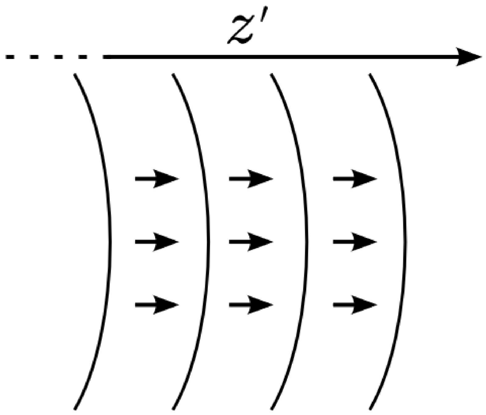


Differentiable Optics



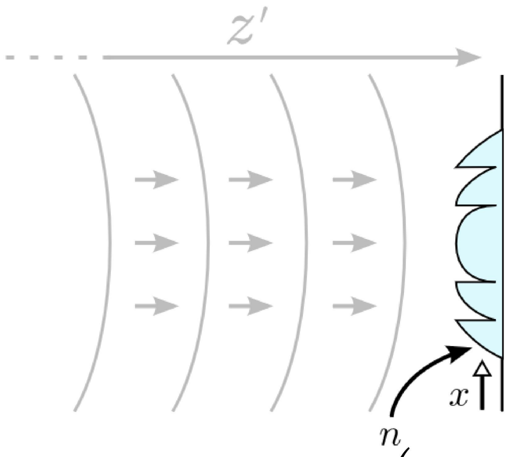
Wave Optics PSF simulator

Spherical wave from point source



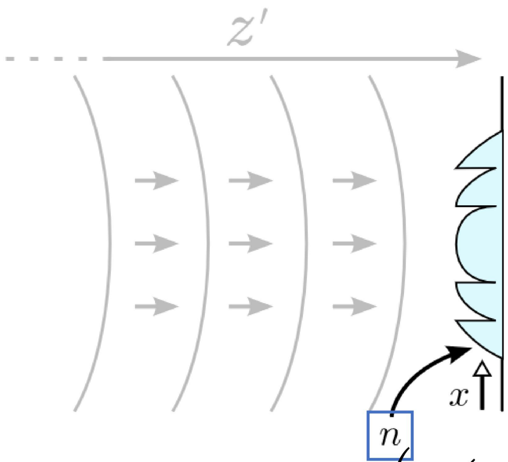
$$exp\left(jk\sqrt{x^2 + z'^2}\right)$$

Phaseshift by optical element



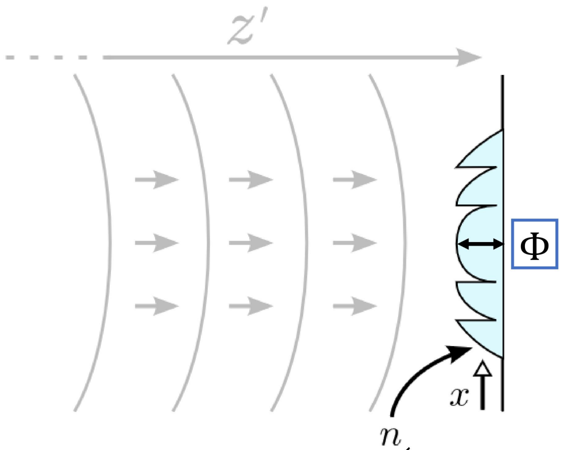
$$U(x) = \exp\left(jk\left(\sqrt{x^2+z'^2} + (n-1)\Phi(\mathrm{x})\right)\right)$$

Phaseshift by optical element



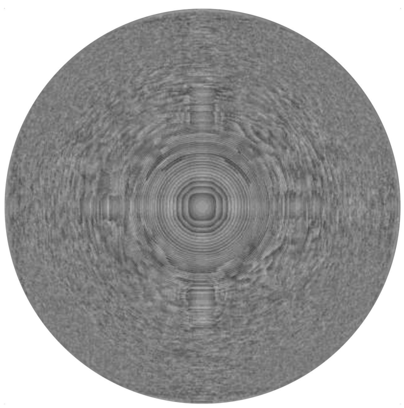
$$U(x) = \exp\left(jk\left(\sqrt{x^2+z'^2} + (\boxed{n}-1)\Phi(\mathrm{x})\right)\right)$$

Phaseshift by optical element



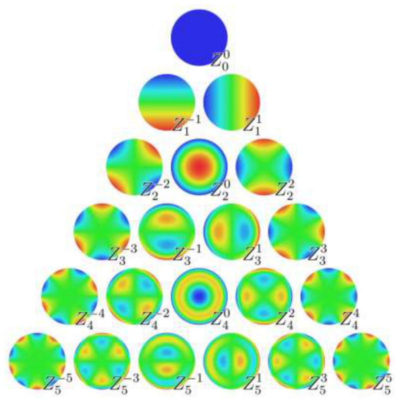
$$U(x) = \exp\left(jk\left(\sqrt{x^2 + z'^2} + (n-1)\boxed{\Phi}(\mathrm{x})\right)\right)$$

Height Map parameterization (diffractive)



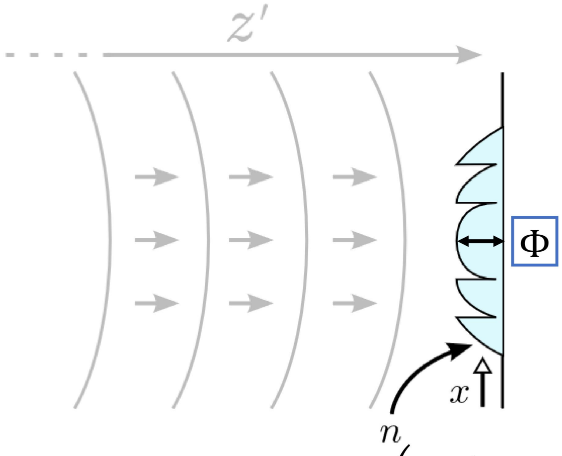
$$\Phi[x] = [[a_{11}, a_{12}, \dots], \dots]$$

Zernike basis parameterization (refractive)



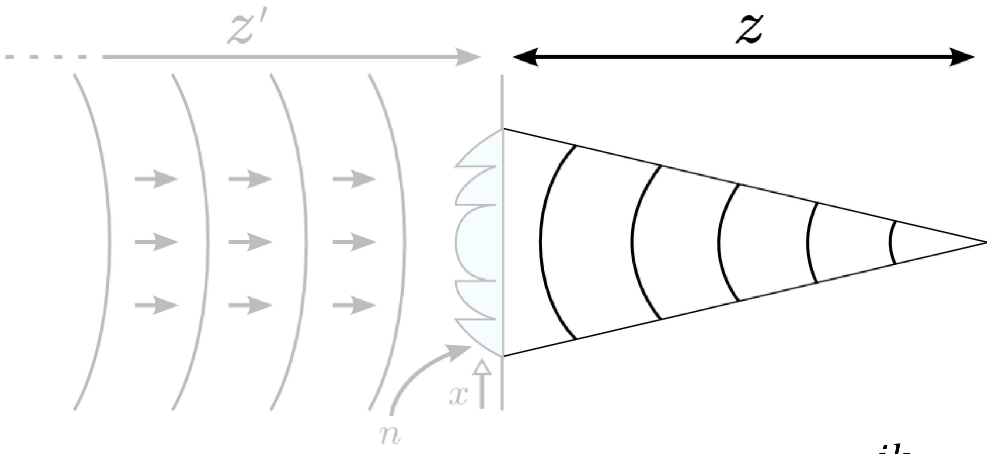
$$\Phi[x] = \sum Z_i^j[x] \cdot a_{ij}$$

Phaseshift by optical element



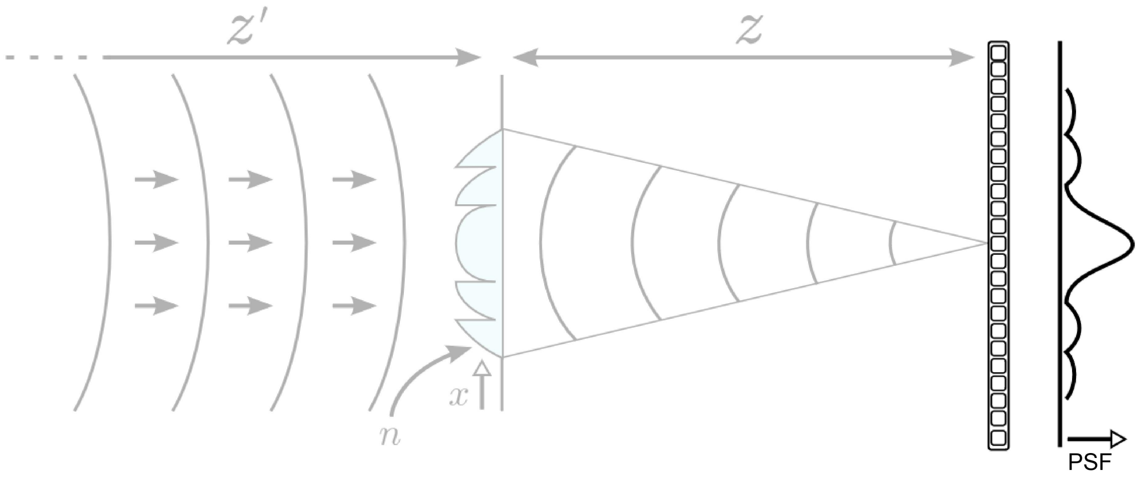
$$U(x) = \exp\left(jk\left(\sqrt{x^2 + z'^2} + (n-1)\boxed{\Phi}(\mathrm{x})\right)\right)$$

Fresnel propagation to sensor



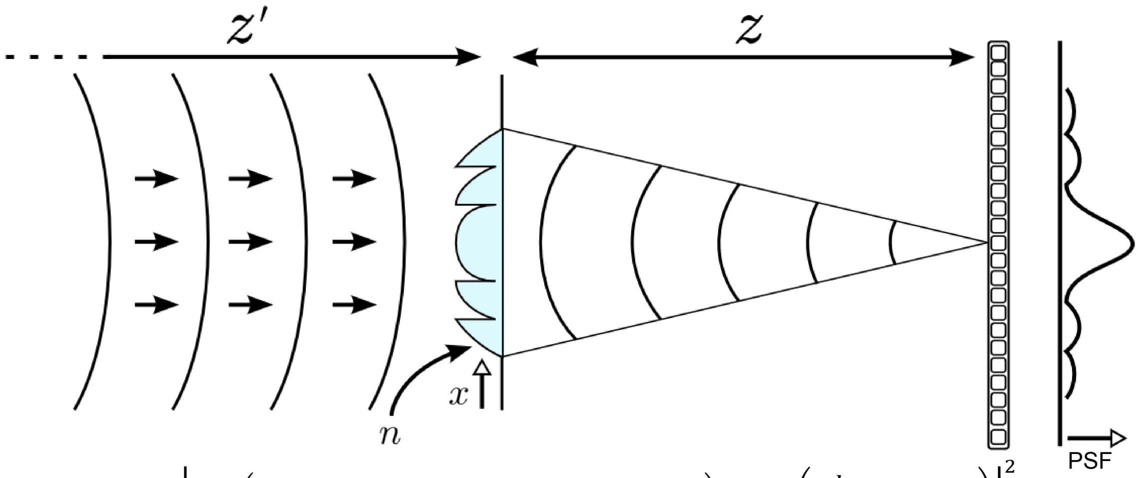
$$U'(x) = U(x) * \exp\left(\frac{jk}{2z}x^2\right)$$

Intensity measurement at sensor



$|U'(x)|^2$

Calculating the PSF

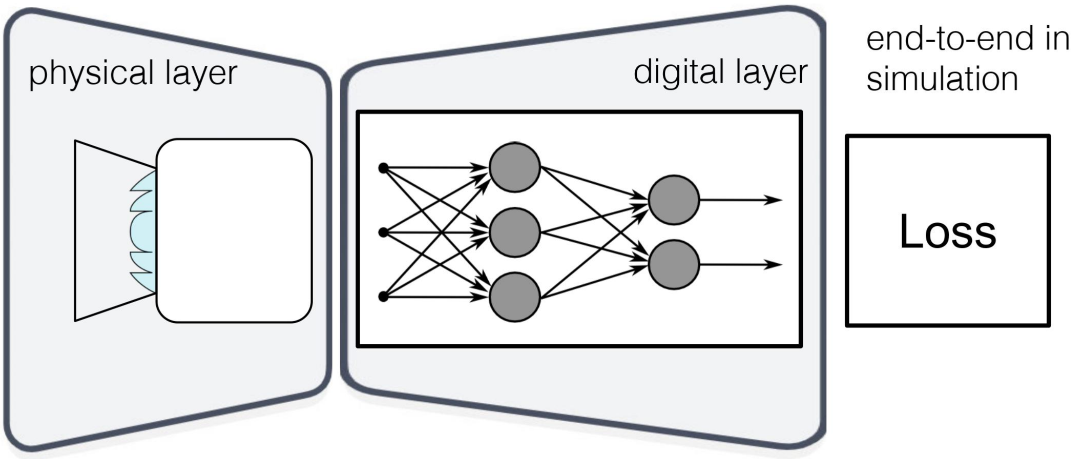


$$\rho_{z',\lambda} = \left| \exp\left(jk \left(\sqrt{x^2 + y^2 + z'^2} + (n-1)\phi(x,y) \right) \right) * \exp\left(j \frac{k}{2z} (x^2 + y^2) \right) \right|^2$$

Deep Optics

Training:

end-to-end in simulation

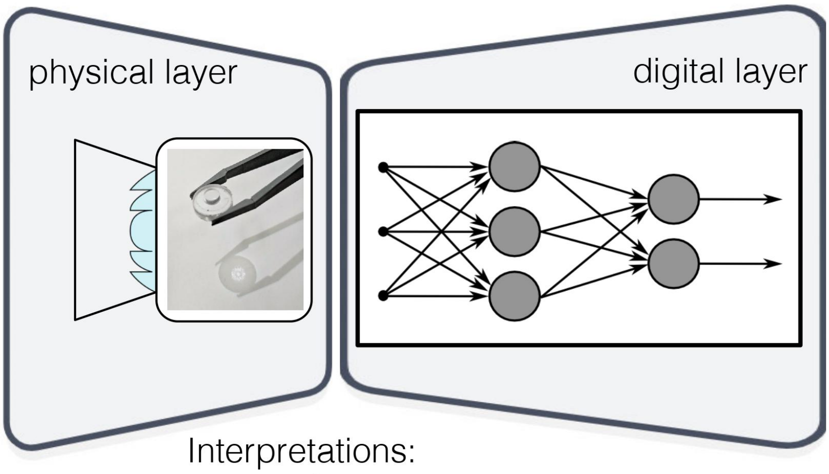


Differentiable Optics

Deep Optics

Inference:

fabricate lens or other physical components, run network



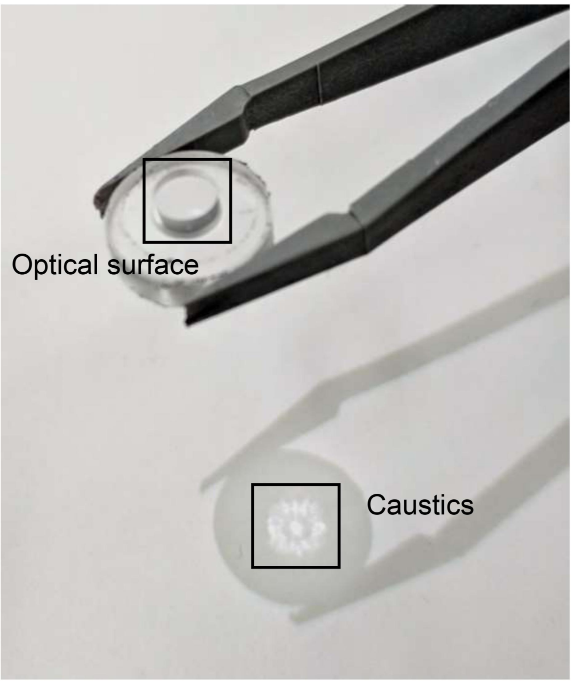
Interpretations:

- Optical encoder, electronic decoder system
- Hybrid optical-electronic neural network

Applications

- Extended Depth of Field
- Monocular Depth Estimation / Depth from Defocus
- High Dynamic Range Imaging
- Video Compressive Sensing
- Computational Microscopy (Will not show examples)

Application 1: Extended Depth of Field (EDOF)



<u>Refractive Achromatic EDOF element</u>

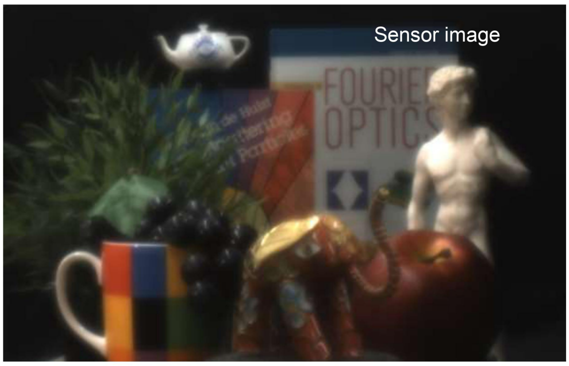
- Polymethyl methacrylate (PMMA)
- 5 mm aperture size
- Sensor distance $35.5mm$
- F-number 7.1
- One active optical surface
- Feature size $3.69\mu m$

- Image source - ACM SIGGRAPH 2018

Application 1: Extended Depth of Field (EDOF)

Test scene

Regular bi-convex lens



Optimized lens

Elephant (0.5m) → Book (2.0m)

- Image source - ACM SIGGRAPH 2018

Application 1: Extended Depth of Field (EDOF)

Test scene

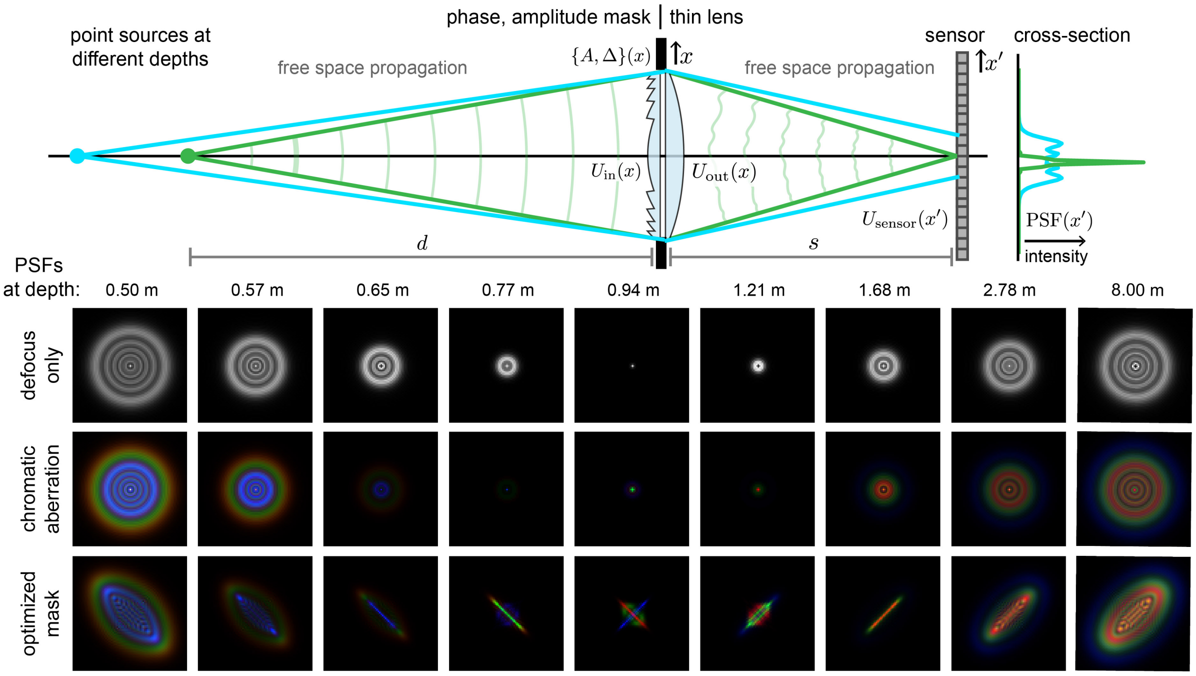
Regular bi-convex lens

Optimized lens

Elephant (0.5m) → Book (2.0m)

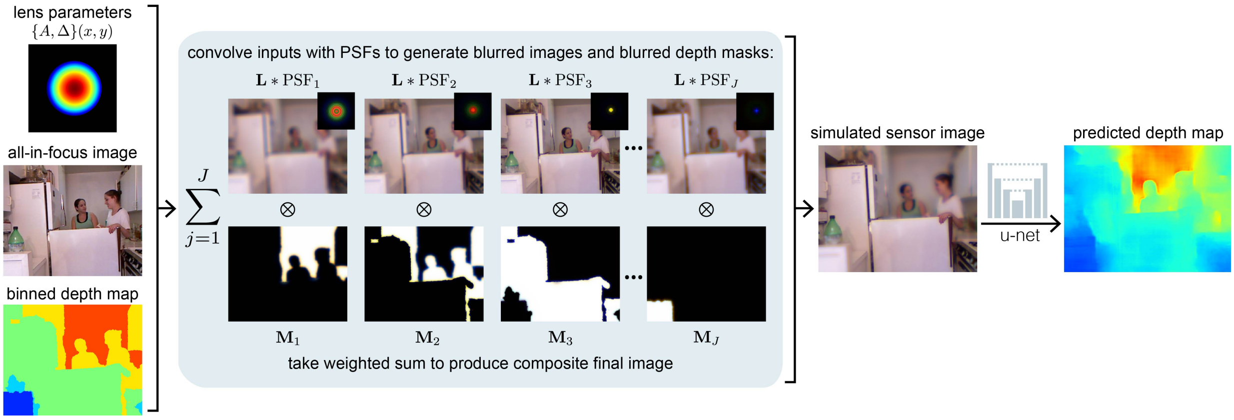
- Image source - ACM SIGGRAPH 2018

Application 2: Monocular Depth Estimation



- Image source - ICCV 2019

Application 2: Monocular Depth Estimation



- Image source - ICCV 2019

In [12]:

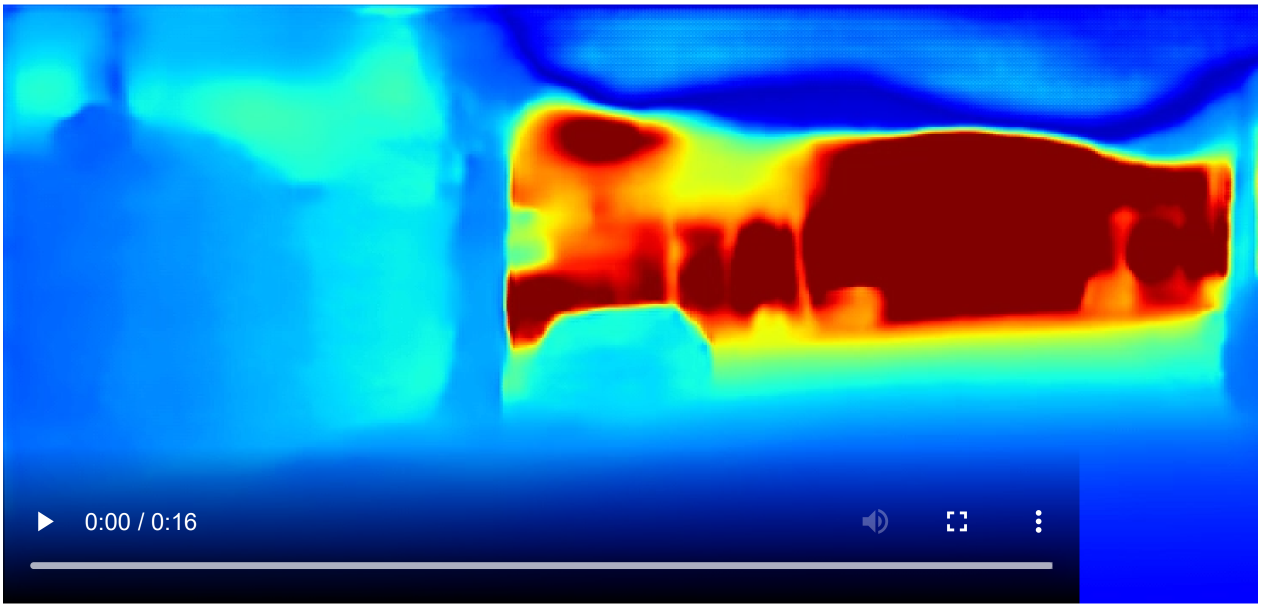
```
from IPython.display import Video
Video("./assets/input_vid_dfd.mp4")
```

Out[12]:

In [10]:

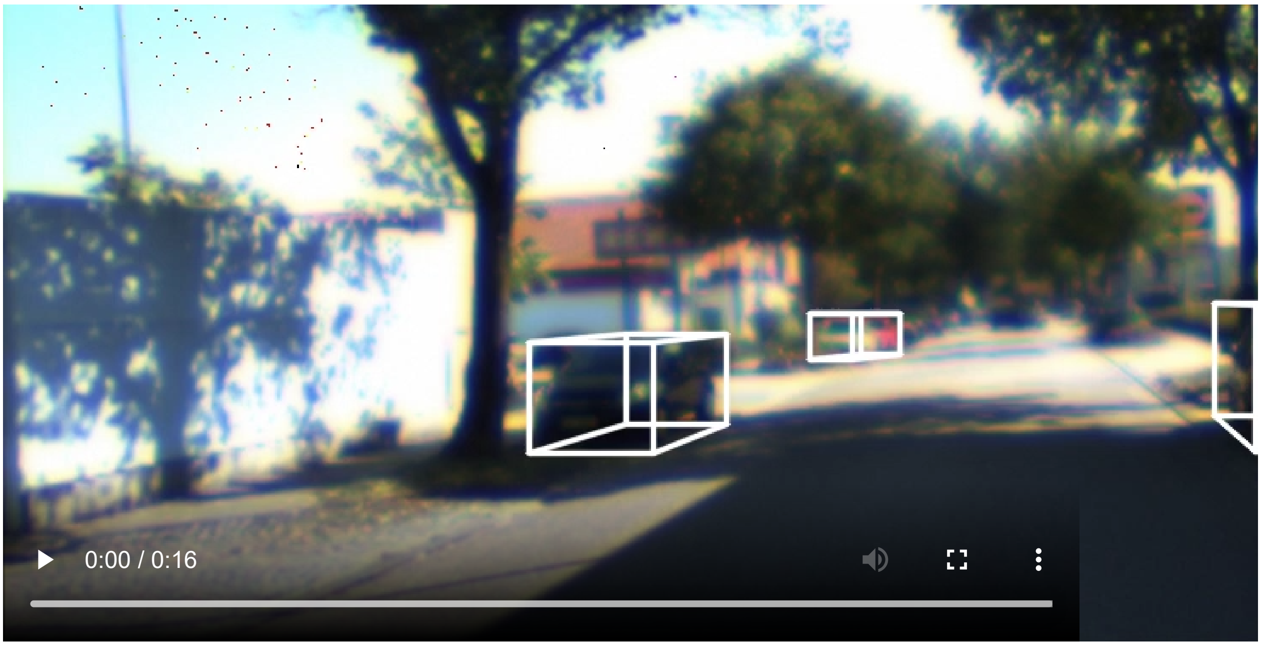
```
Video("./assets/output_vid_dfd.mp4")
```

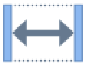
Out[10]:



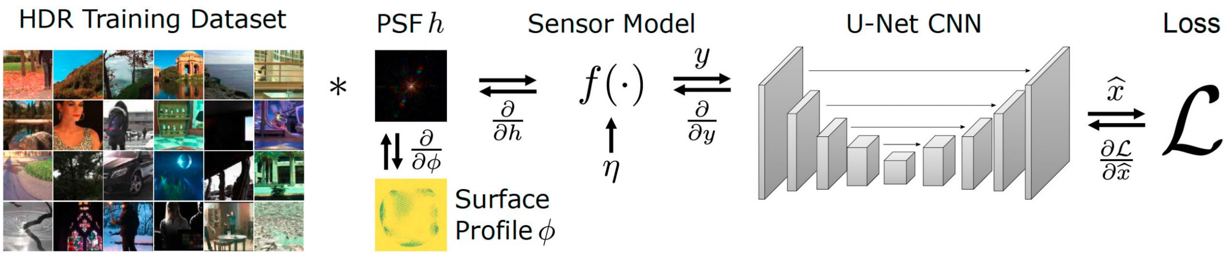
In [13]:
```
Video("./assets/bb_3d.mp4")
```

Out[13]:

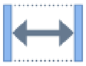




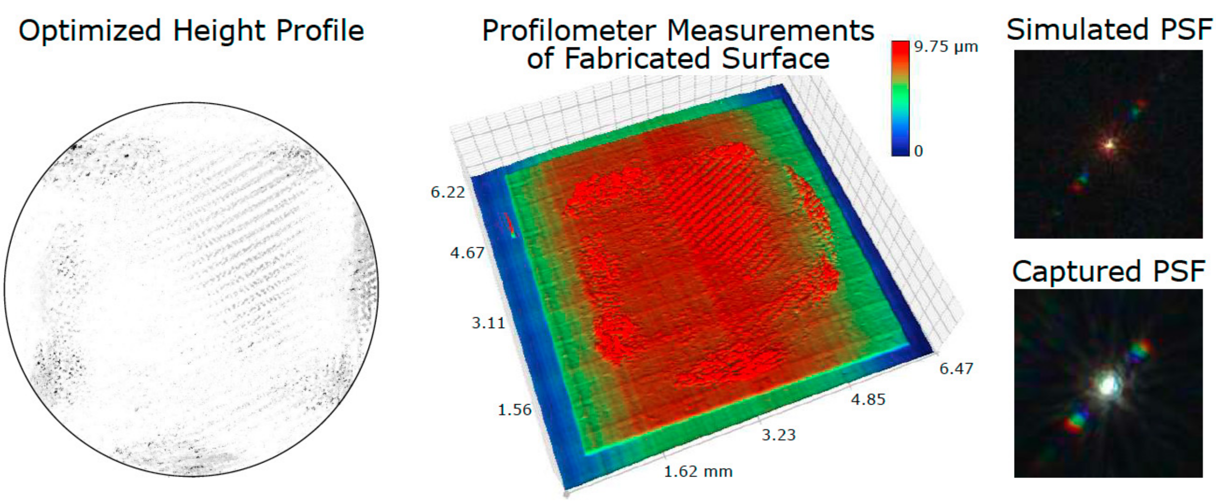
Application 3: High Dynamic Range Imaging



- Image source - CVPR 2020

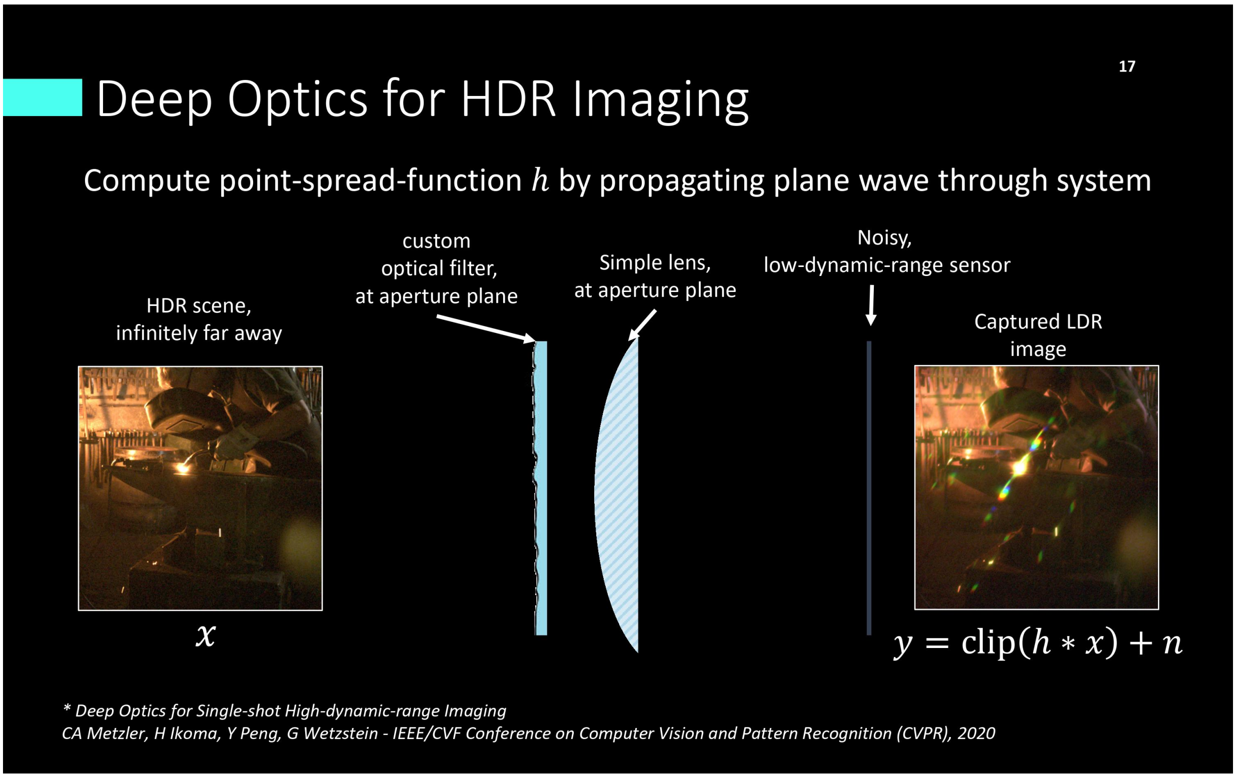


Application 3: High Dynamic Range Imaging



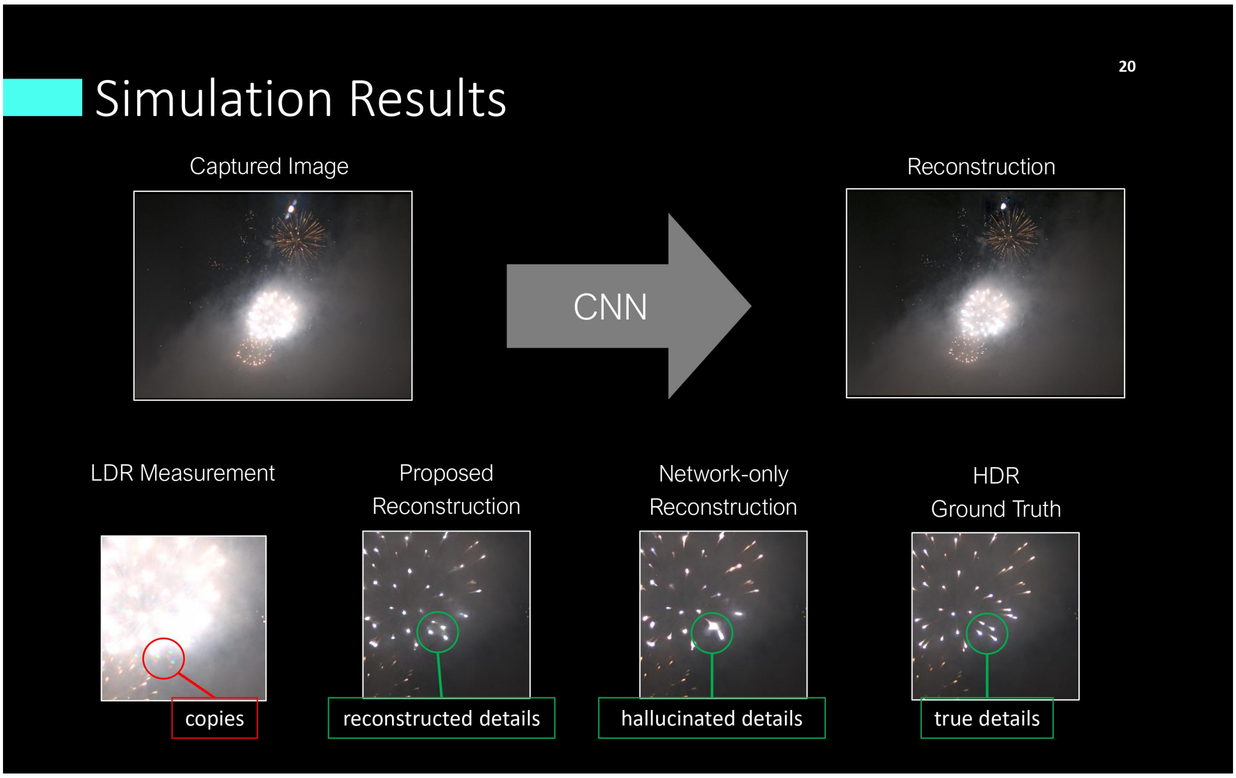
- Image source - CVPR 2020

Application 3: High Dynamic Range Imaging



- Image source - CVPR 2020

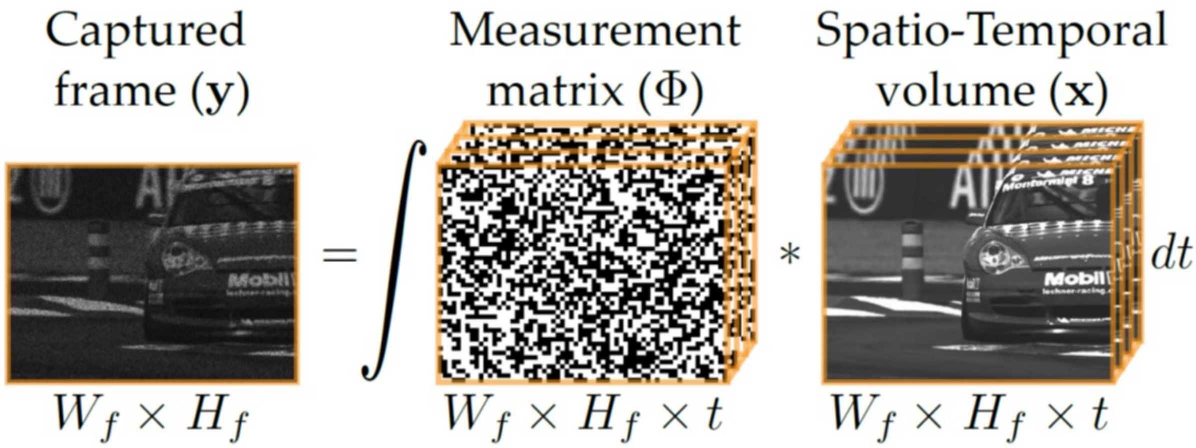
Application 3: High Dynamic Range Imaging



- Image source - CVPR 2020

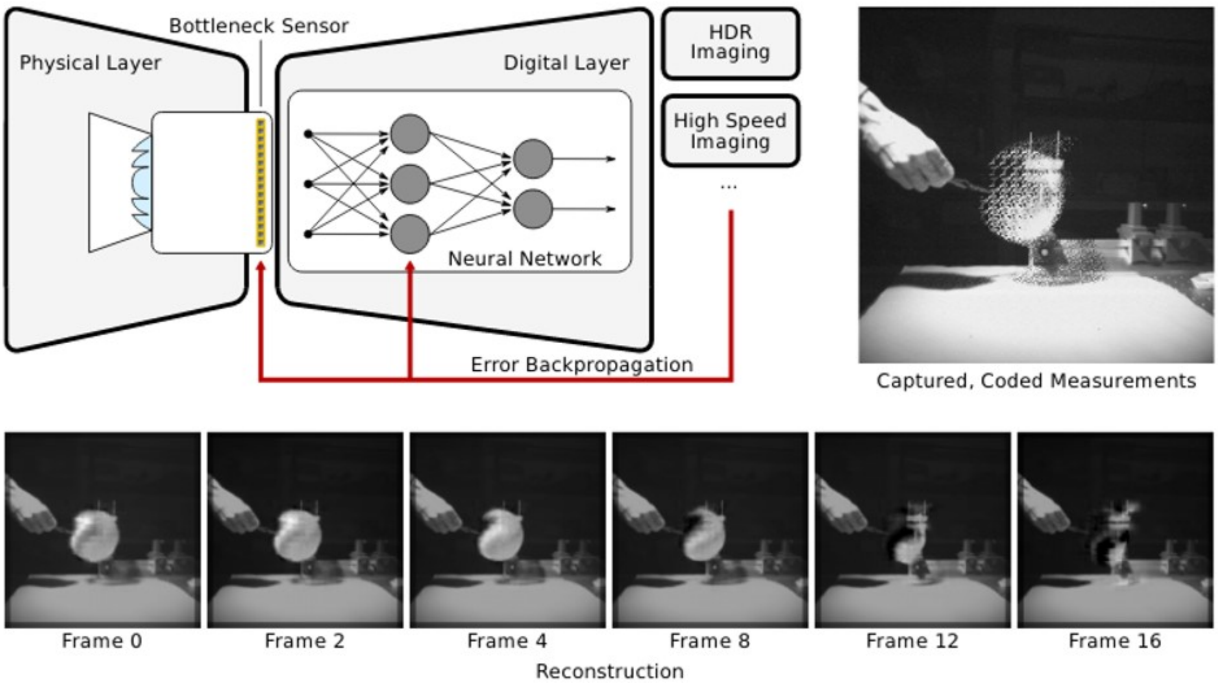
Application 4: Video Compressive Sensing

- Challenge is to learn the binary masks Φ to recover video x from snapshot y



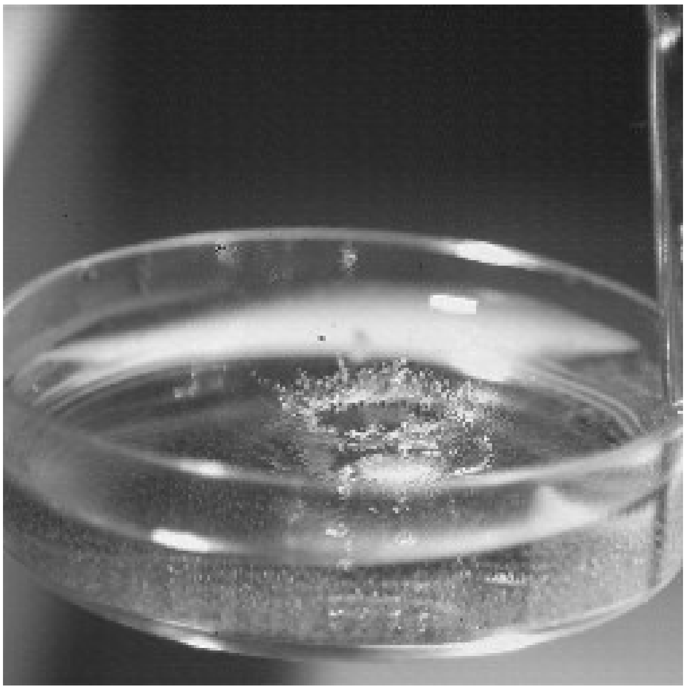
- [Image source - DSP Magazine 2020]

Application 4: Video Compressive Sensing



- [Image source - ICCP 2020]

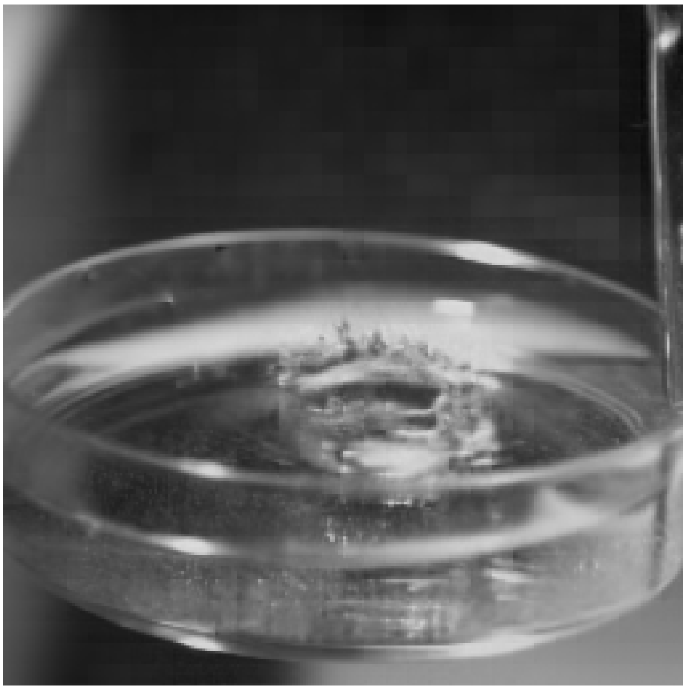
Application 4: Video Compressive Sensing



- Image source - ICCP 2020

Application 4: Video Compressive Sensing

- 64 frames recovered from 4 measurements (16 frames/capture)



- Image source - ICCP 2020

Available resources online

- https://github.com/computational-imaging/opticalCNN
- https://github.com/vsitzmann/deepoptics
- https://github.com/computational-imaging/DeepOpticsHDR
- https://github.com/EliasNehme/DeepSTORM3D

- https://github.com/computational-imaging/DepthFromDefocusWithLearnedOptics
- etc.

Recommended Videos

Warning!

- These videos do not replace the lectures and tutorials.
- Please use these to get a better understanding of the material, and not as an alternative to the written material.

Video By Subject

- End-to-end Optimization of Optics and Image Processing - Vincent Sitzmann
- Neural Sensors - J.N.P Martel
- High Dynamic Range Imaging - Christopher Metzler
- 3D Single Molecule Localization Microscopy - Elias Nehme
- Towards Neural Imaging & Signal Processing - Gordon Wetzstein

Credits

- End-to-end optimization of optics and image processing - Vincent Sitzmann
- ACM SIGGRAPH 2020 Courses - Yifan (Evan) Peng, Ashok Veeraraghavan, Wolfgang Heidrich, Gordon Wetzstein
- Stanford EE367/CS448I (Computational Imaging and Display) - Gordon Wetzstein
- IEEE SPACE Webinar - IEEE Computational Imaging TC
- Research papers:
  - End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging
  - Deep Optics for Monocular Depth Estimation and 3D Object Detection
  - Deep Optics for Single-shot High-dynamic-range Imaging
  - Neural Sensors: Learning Pixel Exposures for HDR Imaging and Video Compressive Sensing With Programmable Sensors
  - Convolutional neural networks that teach microscopes how to image
  - DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning
  - Depth From Defocus With Learned Optics
  - etc.
- Icons from Icon8.com - https://icons8.com