CS-ToF: High-resolution Compressive time-of-flight imaging

Fengqiang Li, Chia-kai Yeh, Kuan He, Oliver Cossairt (Northwestern University)
Huaijin Chen, Adithya Pediredla, Ashok Veeraraghavan (Rice University)

fengqiang.li@u.northwestern.edu

Time of flight

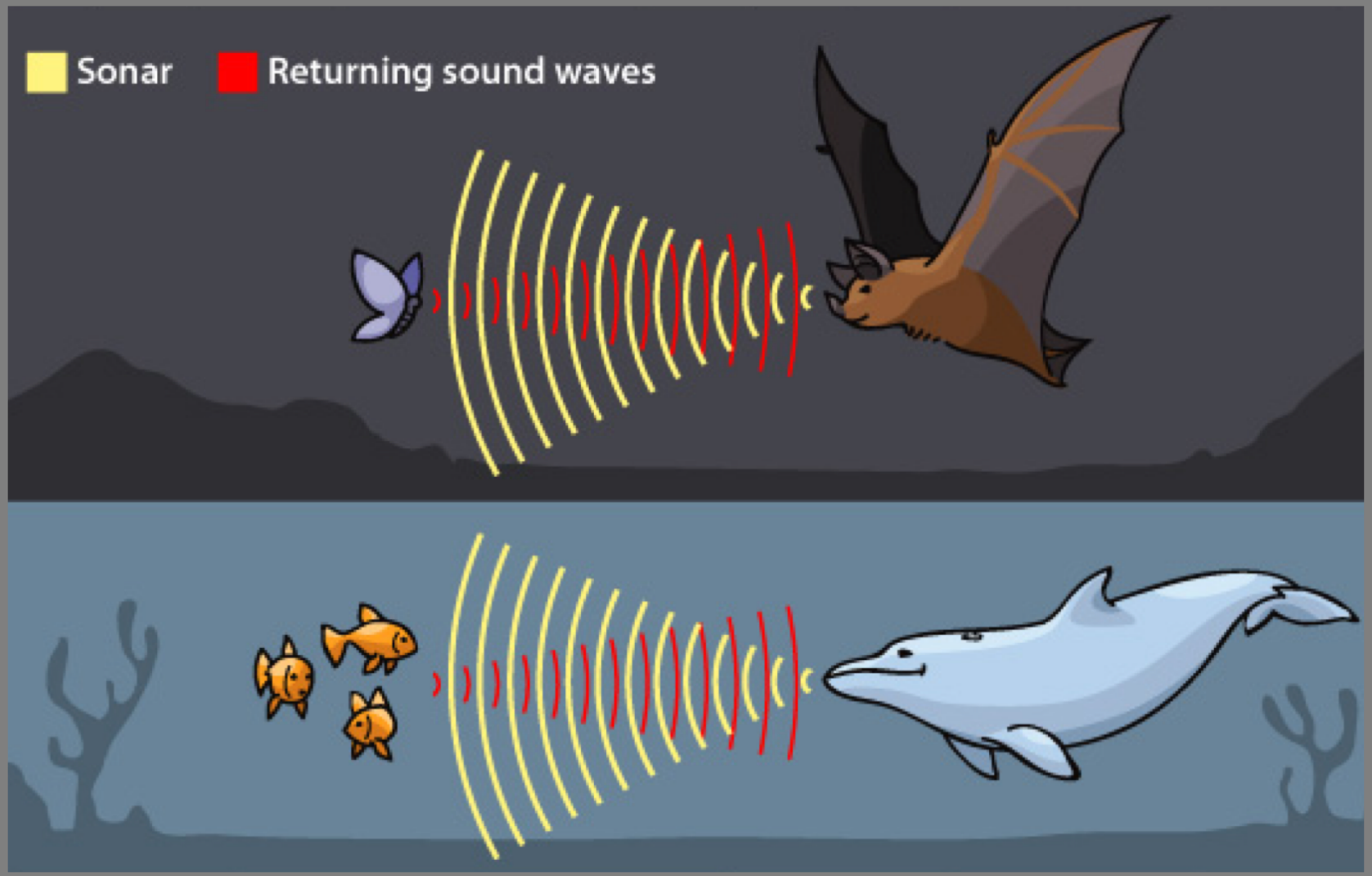
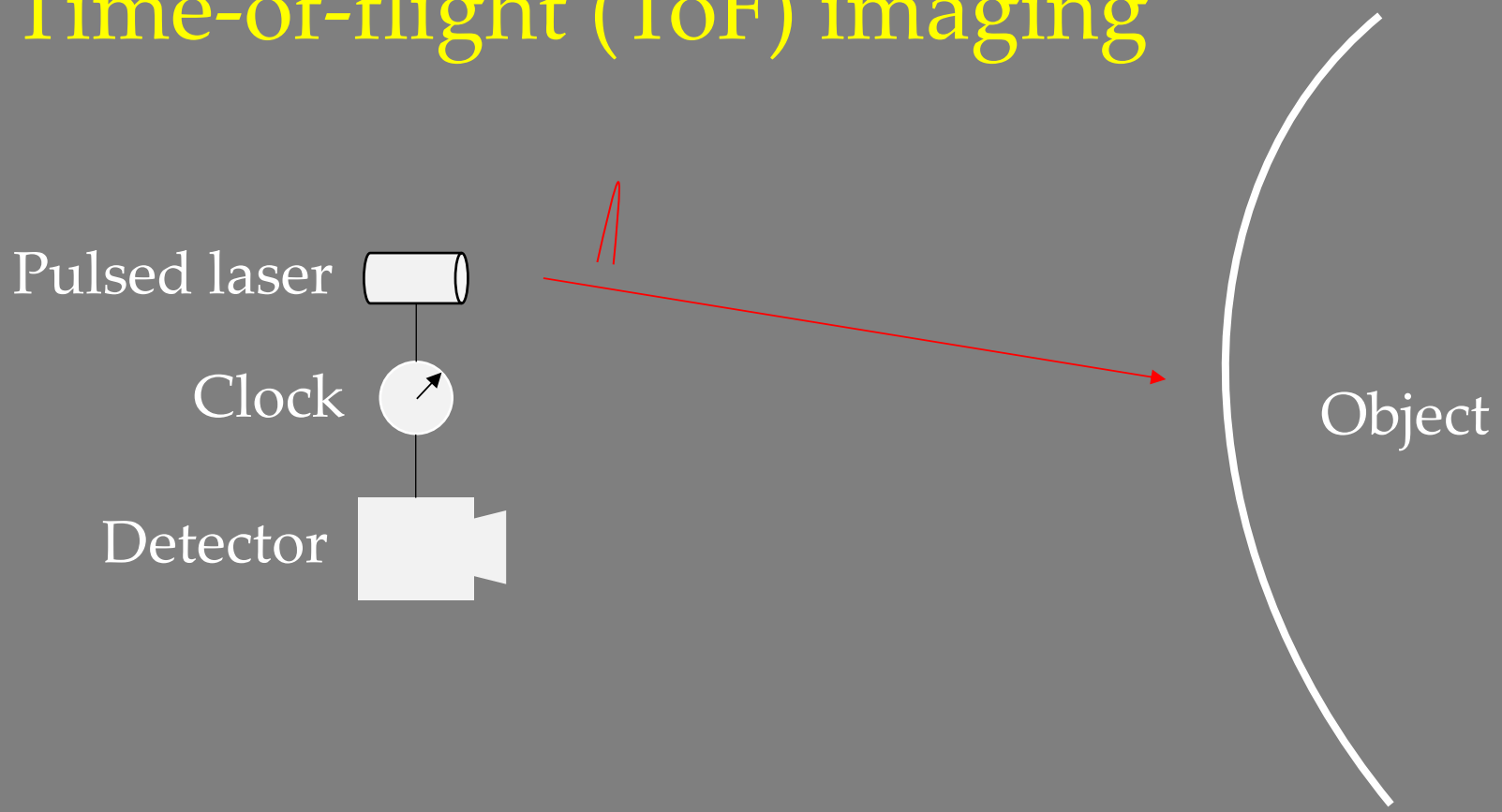


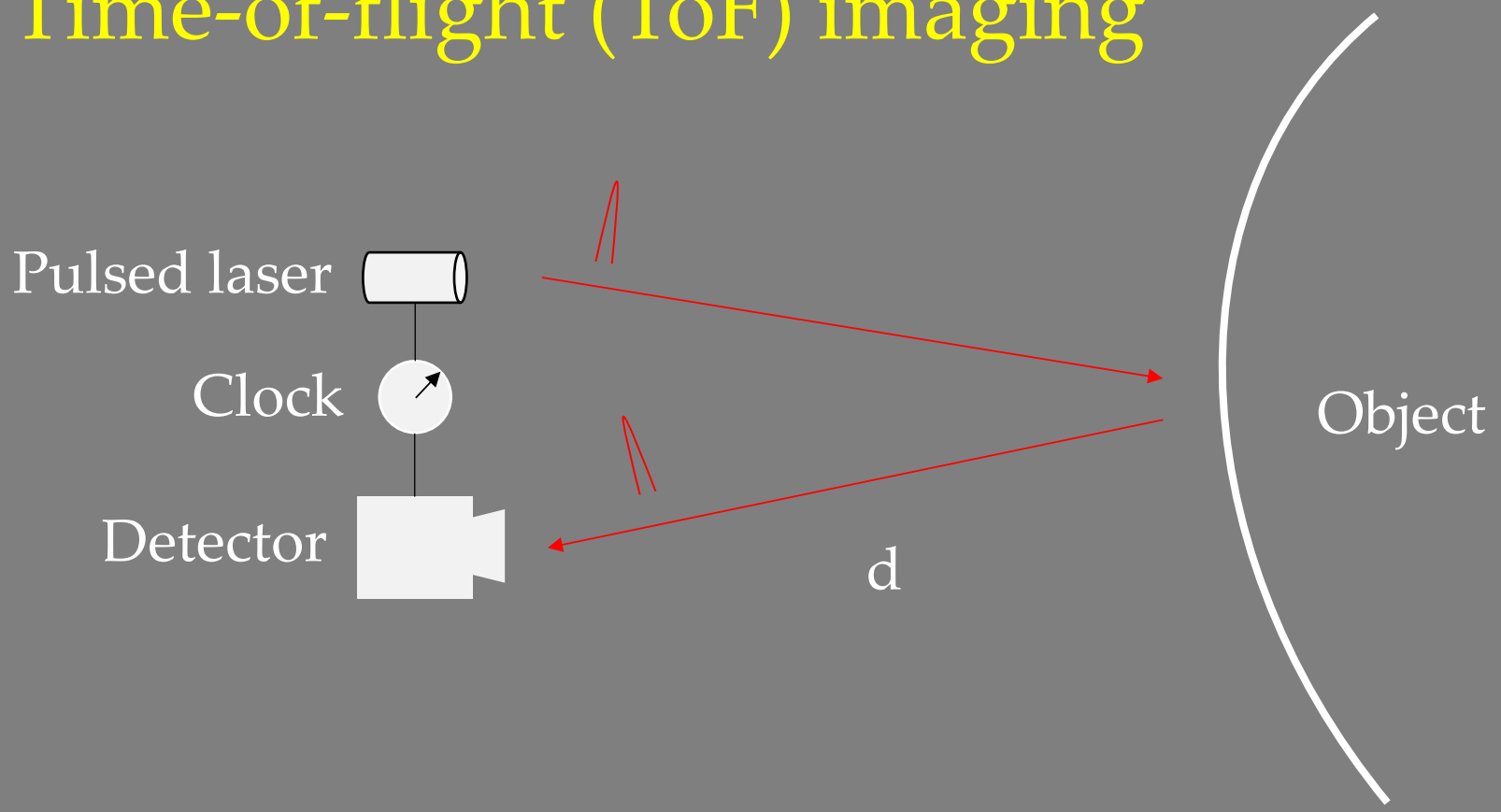
Image credit: Donald Griffin

Time-of-flight (ToF) imaging



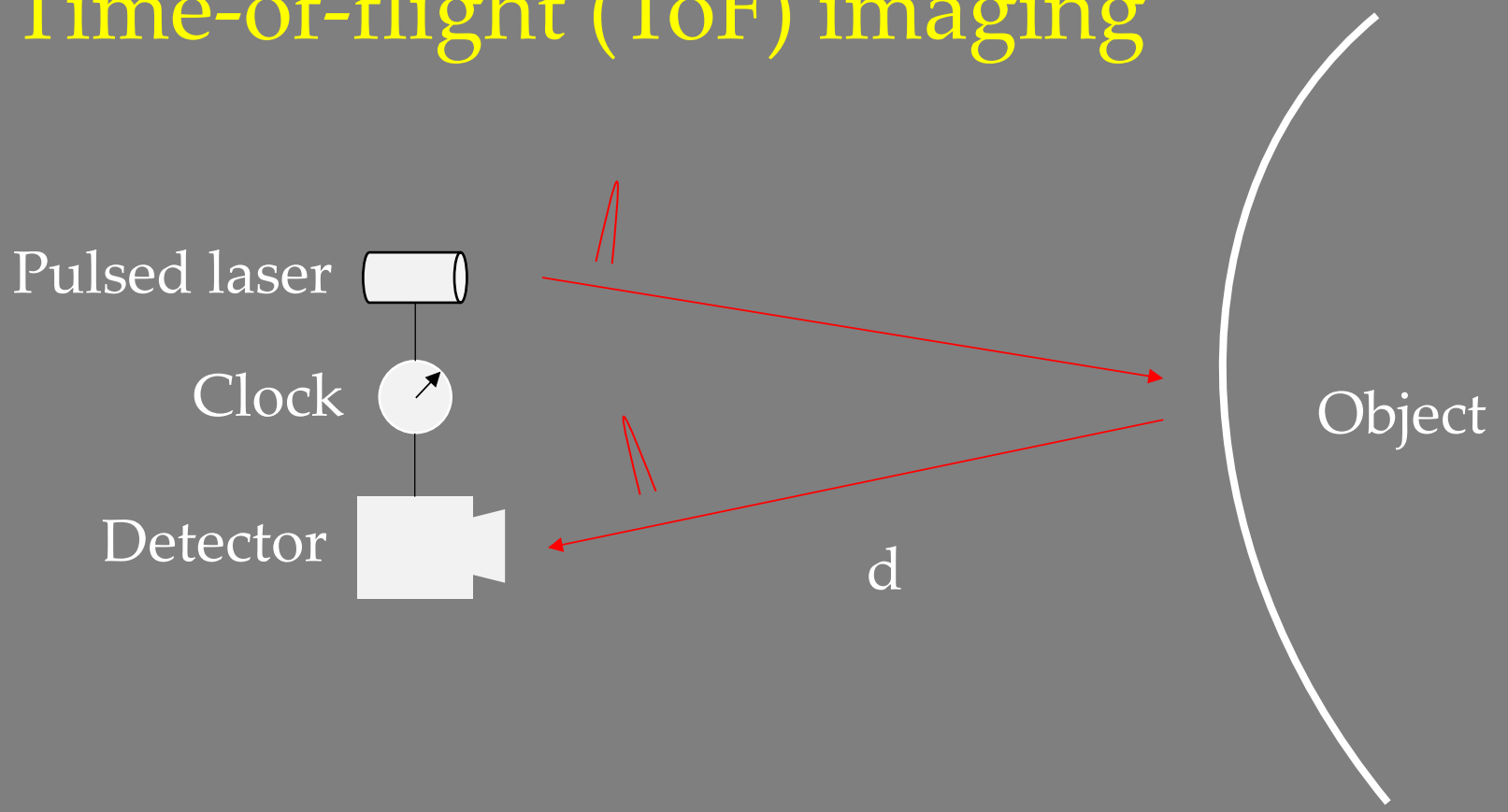
Goyer et al., *Bulletin of the American Meteorological Society,* 1963

Time-of-flight (ToF) imaging



Goyer et al., *Bulletin of the American Meteorological Society,* 1963

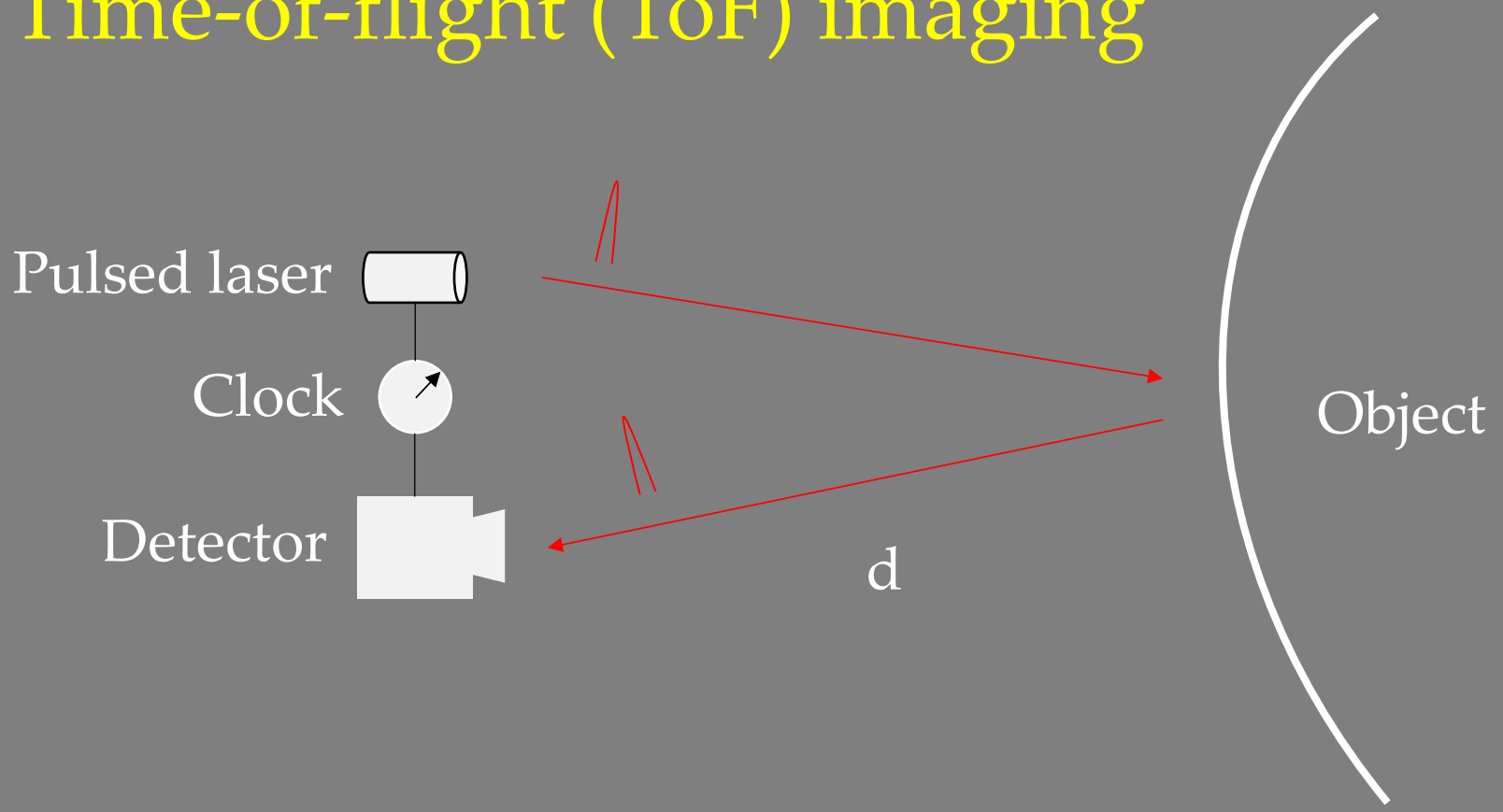
Time-of-flight (ToF) imaging



$$d = \frac{1}{2} \cdot c \cdot t$$

Goyer et al., *Bulletin of the American Meteorological Society,* 1963

Time-of-flight (ToF) imaging



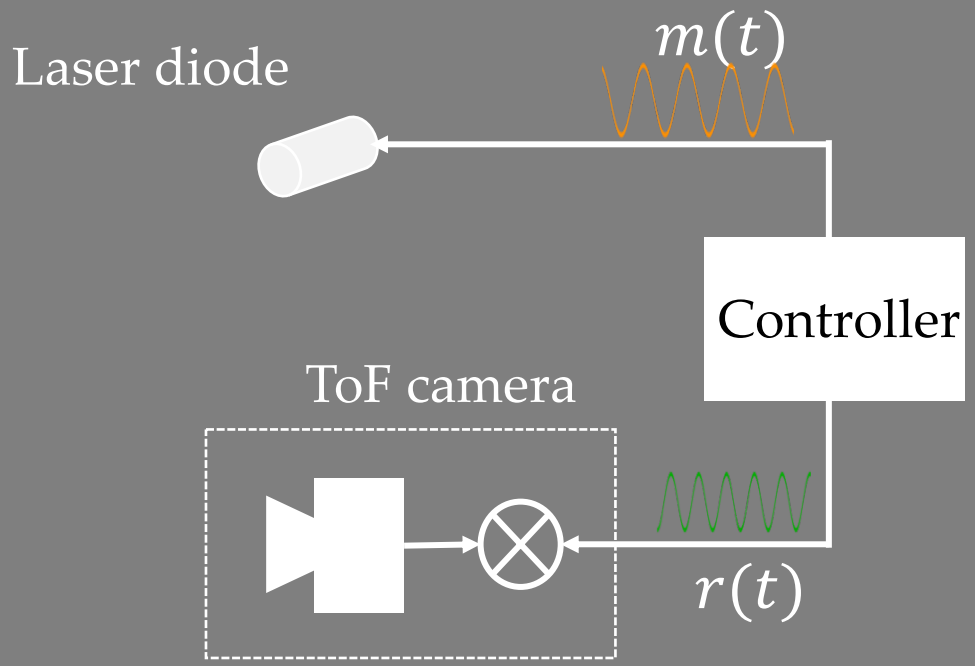
$$d = \frac{1}{2} \cdot c \cdot t$$

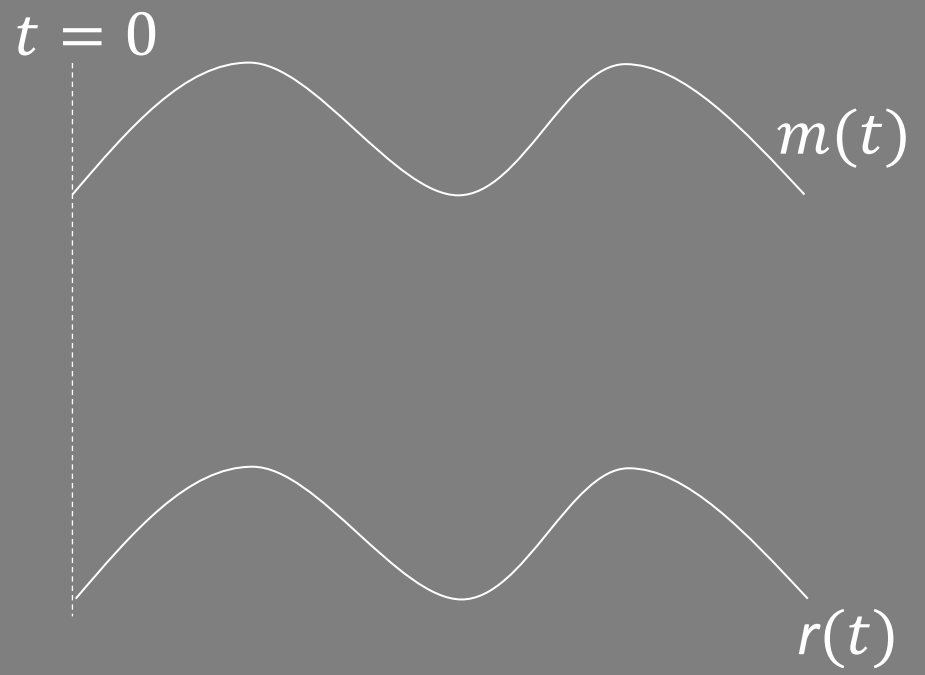
Pulsed light based ToF

- Raster scanning

Goyer et al., *Bulletin of the American Meteorological Society*, 1963

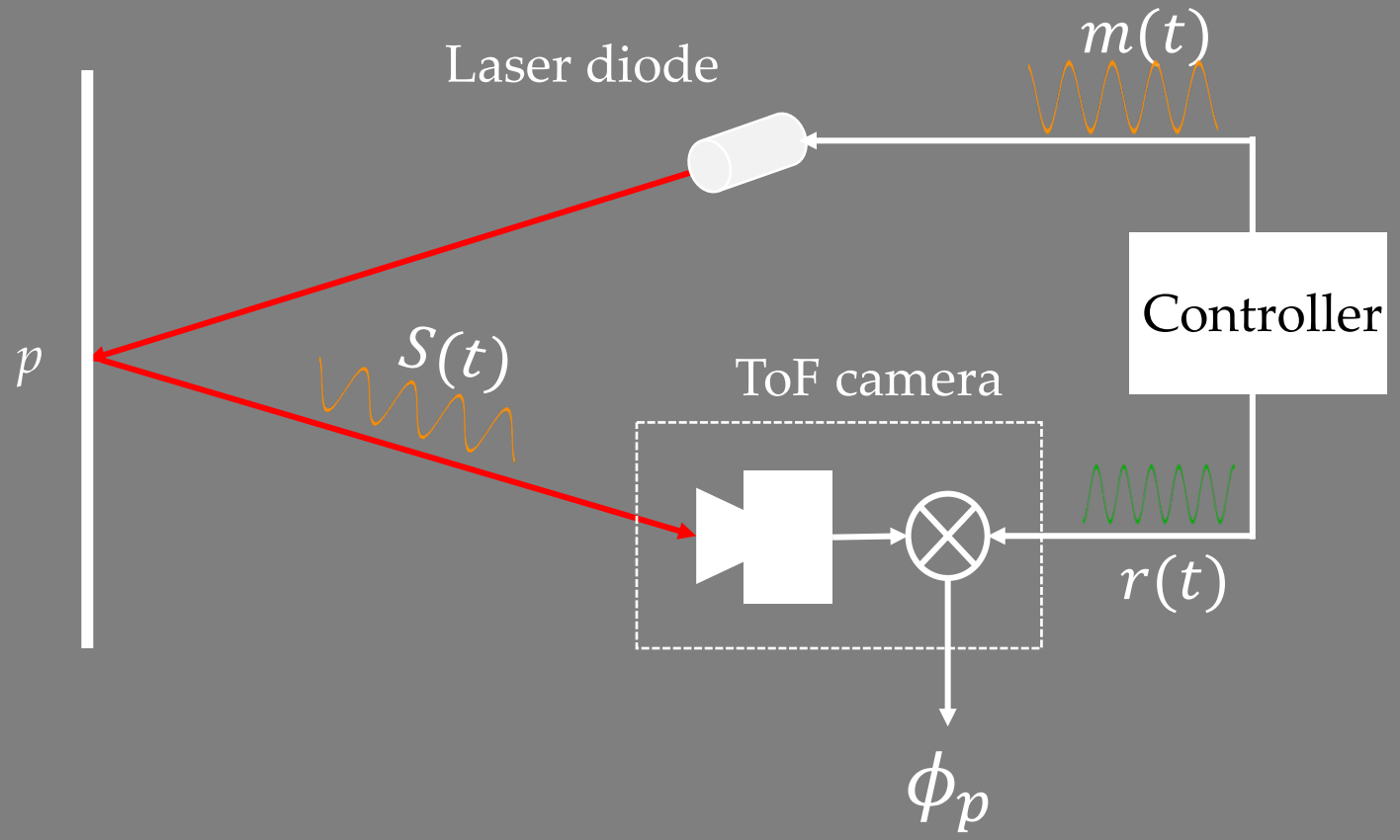
Continuous-wave ToF

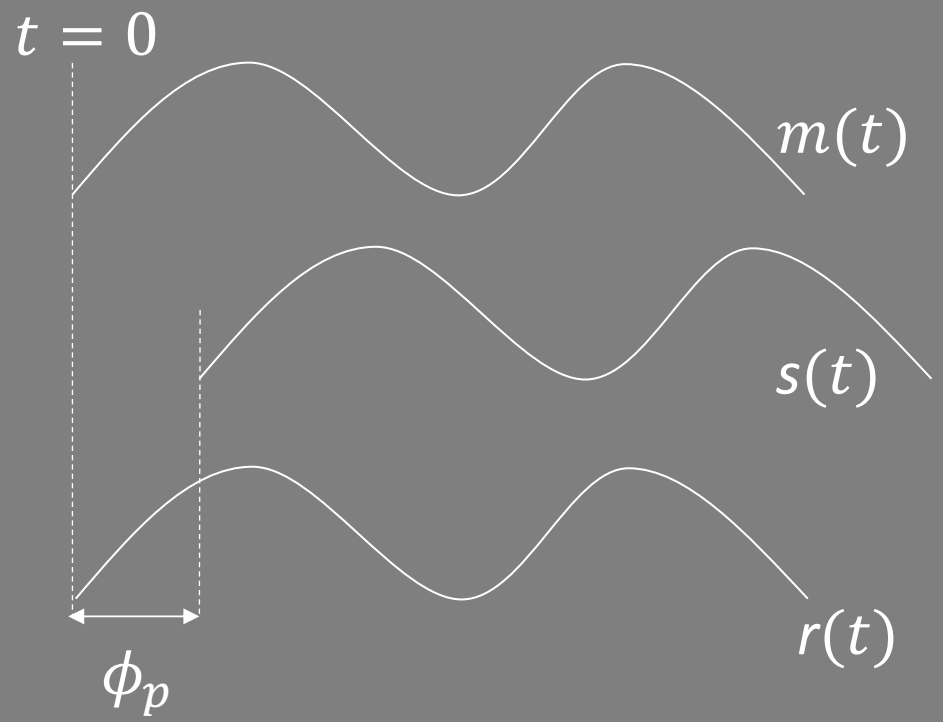




Schwarte et al., SPIE Proceeding, 1997

Continuous-wave ToF





Schwarte et al., SPIE Proceeding, 1997

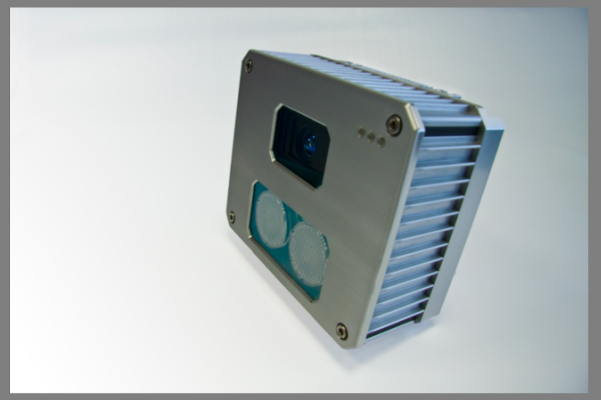

JOBY

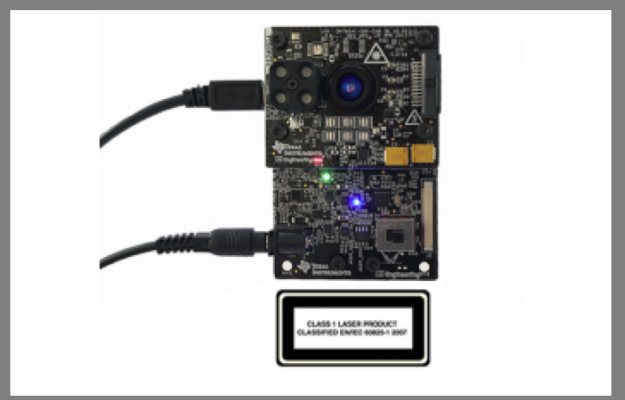

PMD
TECHNOLOGIES


MESA


XBOX


Panasonic
EKL3104




CLASS 1 LASER PRODUCT

Autonomous cars

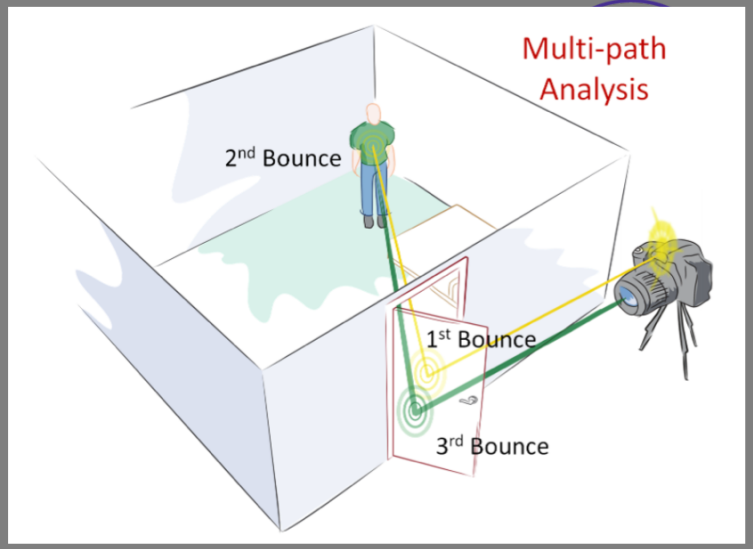
HCI

Robotics

VR/AR

Autonomous cars

HCI

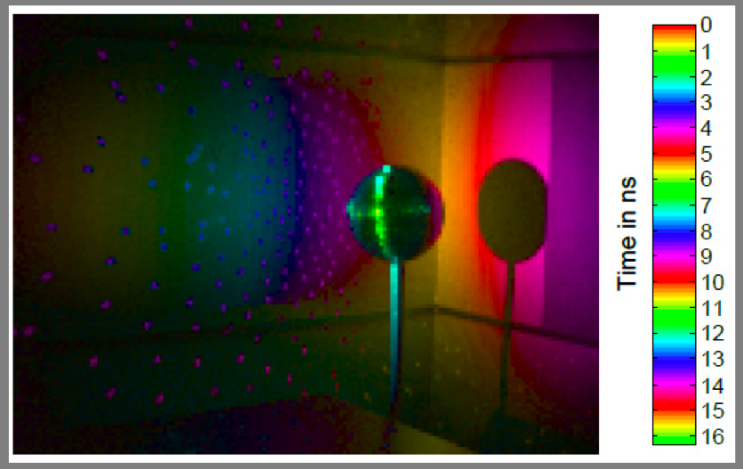


Non-Line-of-Sight imaging

Robotics

Project Tango

VR/AR



Transient imaging

Low spatial resolution

- 640×480 pixels (0.3 mega) vs 120-megapixel CMOS

Low spatial resolution

- 640×480 pixels (0.3 mega) vs 120-megapixel CMOS
- Extra circuits for each pixel
- Limited wafer size

Previous work

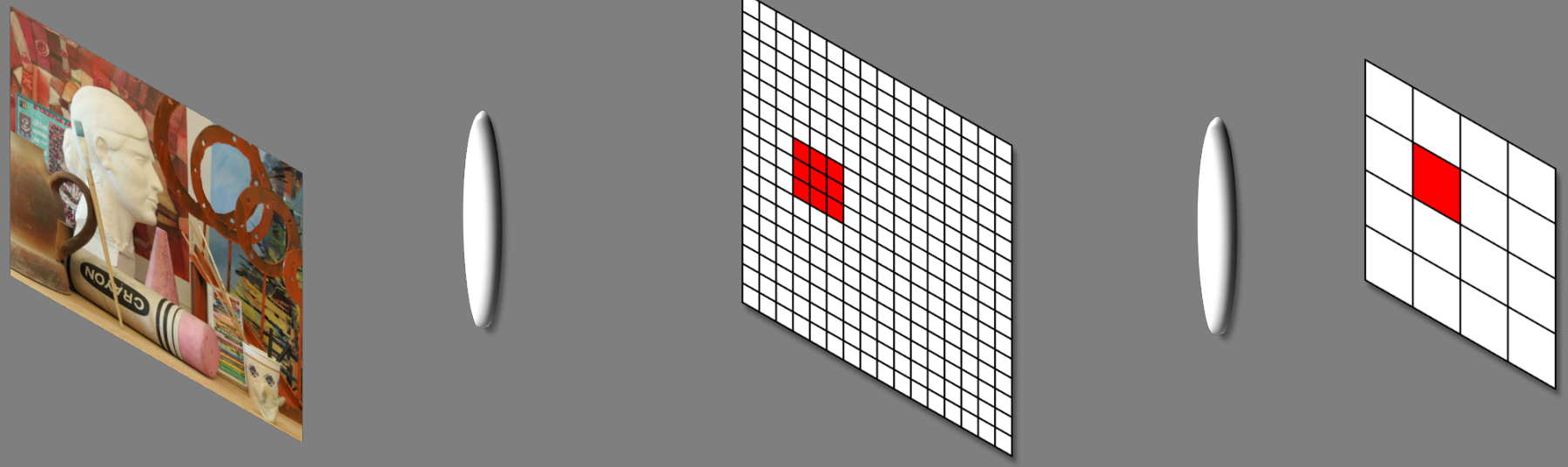
Work directly on ToF camera output:

- Edge guidance
- Defocus debluring

Fuse ToF output with a second camera:

- RGB camera
- Stereo, Photometric stereo, Shape from polarization

Optical multiplexing



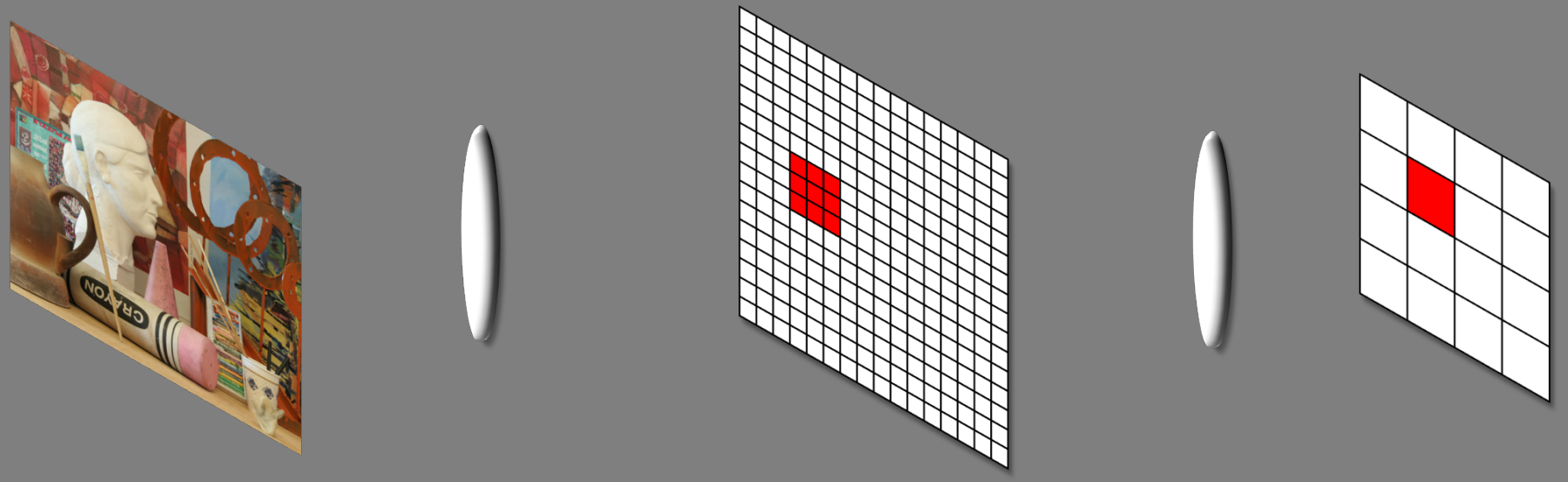
Objects

Spatial light modulator-SLM
(High resolution)

ToF camera
(low resolution)

- **Multiple Pixels (e.g. 3×3)** on spatial light modulator projected on **One Pixel** of ToF camera

Motivation for optical multiplexing



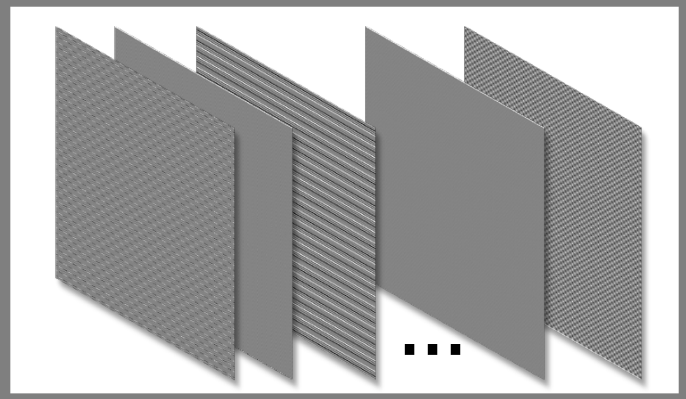
- Higher resolution compared to algorithm based method
- No need to fusion with different image modalities
- Resolution is dependent on SLM

Compressive sensing

Objects

Spatial light modulator

ToF camera

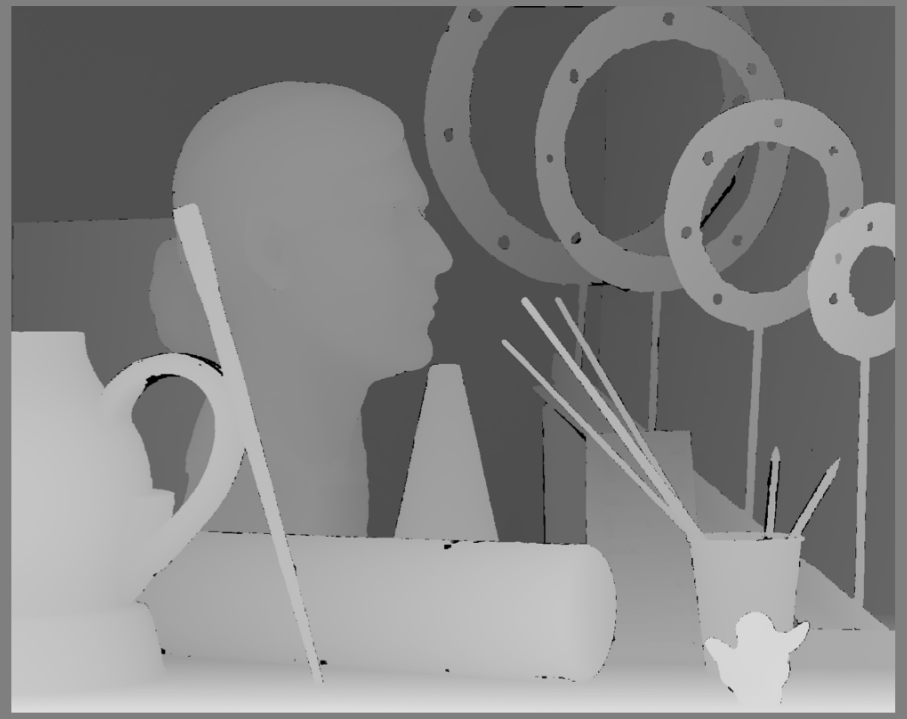


Pattern on modulator (M)

ToF output (y)

Phasor

Intensity (a)



depth (d)

- ToF output: intensity (a) and depth (d)
- Phasor: $a \circ e^{i\frac{4\pi f}{c}d} = a \circ e^{i\phi}$

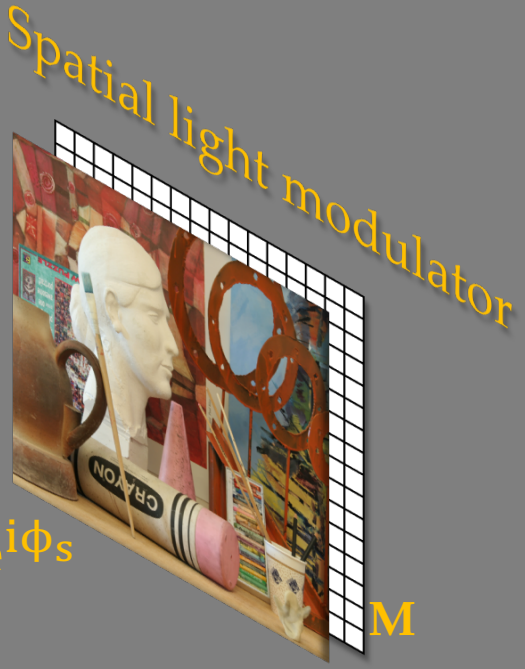
Imaging forward model

$\mathbf{x} = \mathbf{a_s} \circ e^{i\phi_s}$

- High resolution scene projection on modulator (**x**)

x

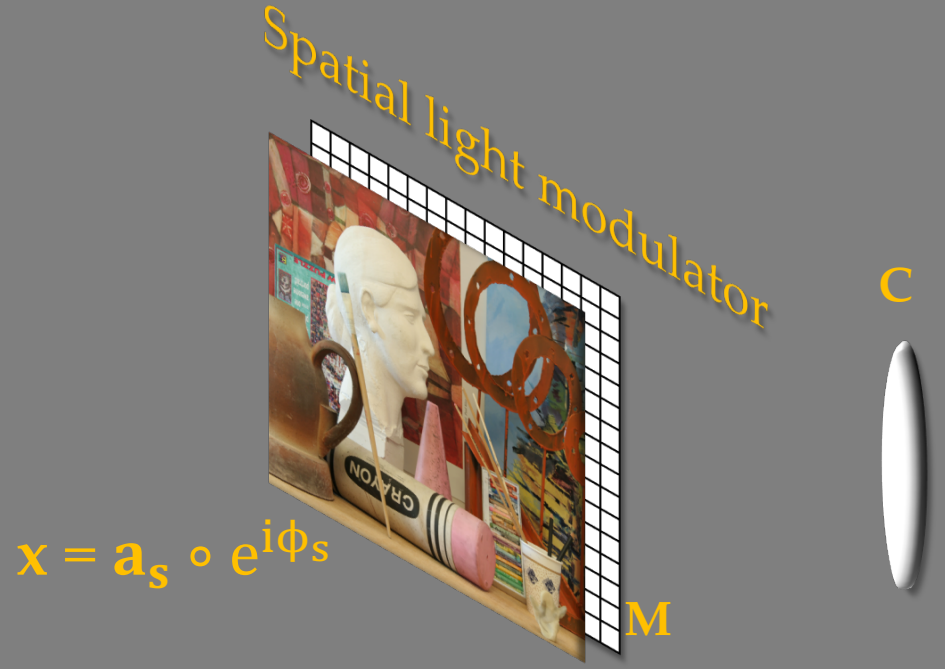
Imaging forward model



- High resolution scene projection on modulator (**x**)
- Spatial light modulator pattern (**M**)

Mx

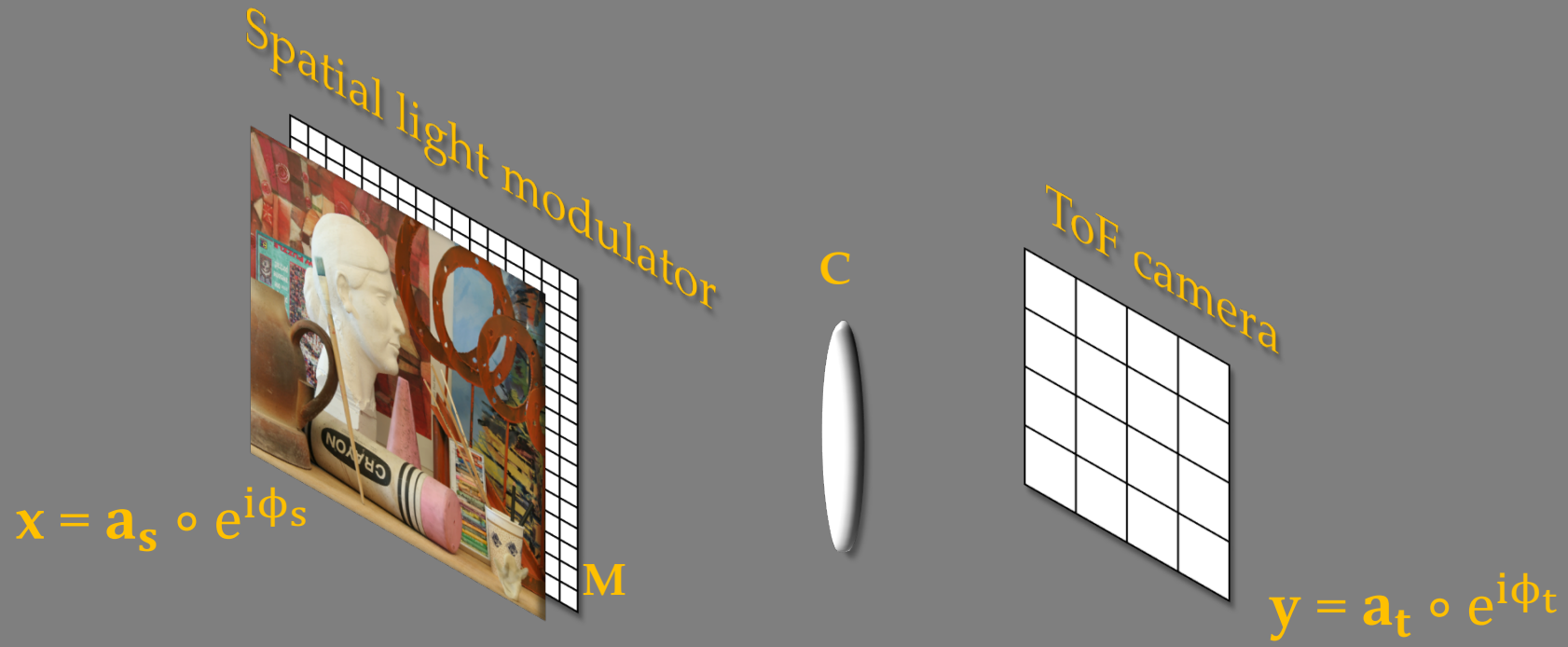
Imaging forward model



- High resolution scene projection on modulator (**x**)
- Spatial light modulator pattern (**M**)
- Translation matrix from modulator to ToF camera (**C**)

CMx

Imaging forward model



- High resolution scene projection on modulator (**x**)
- Spatial light modulator pattern (**M**)
- Translation matrix from modulator to ToF camera (**C**)
- ToF camera output (**y**)

$$y = CMx$$

Imaging forward model

- ToF Output y_i with modulation pattern M_i

$$\begin{bmatrix} a_{t1}e^{i\phi_{t1}} \\ a_{t2}e^{i\phi_{t2}} \\ a_{t3}e^{i\phi_{t3}} \\ .. \\ a_{tN}e^{i\phi_{tN}} \end{bmatrix} = \begin{bmatrix} \mathbf{A_1}a_{s1}e^{i\phi_{s1}} \\ \mathbf{A_2}a_{s2}e^{i\phi_{s2}} \\ \mathbf{A_3}a_{s3}e^{i\phi_{s3}} \\ .. \\ \mathbf{A_N}a_{sN}e^{i\phi_{sN}} \end{bmatrix} = \begin{bmatrix} \mathbf{A_1} \\ \mathbf{A_2} \\ \mathbf{A_3} \\ .. \\ \mathbf{A_N} \end{bmatrix} a_s e^{i\phi_s} \quad \Rightarrow \quad \begin{bmatrix} \mathbf{y_1} \\ \mathbf{y_2} \\ \mathbf{y_3} \\ .. \\ \mathbf{y_N} \end{bmatrix} = \begin{bmatrix} \mathbf{A_1} \\ \mathbf{A_2} \\ \mathbf{A_3} \\ .. \\ \mathbf{A_N} \end{bmatrix} \mathbf{x}$$

$$\boldsymbol{y = CMx = Ax}$$

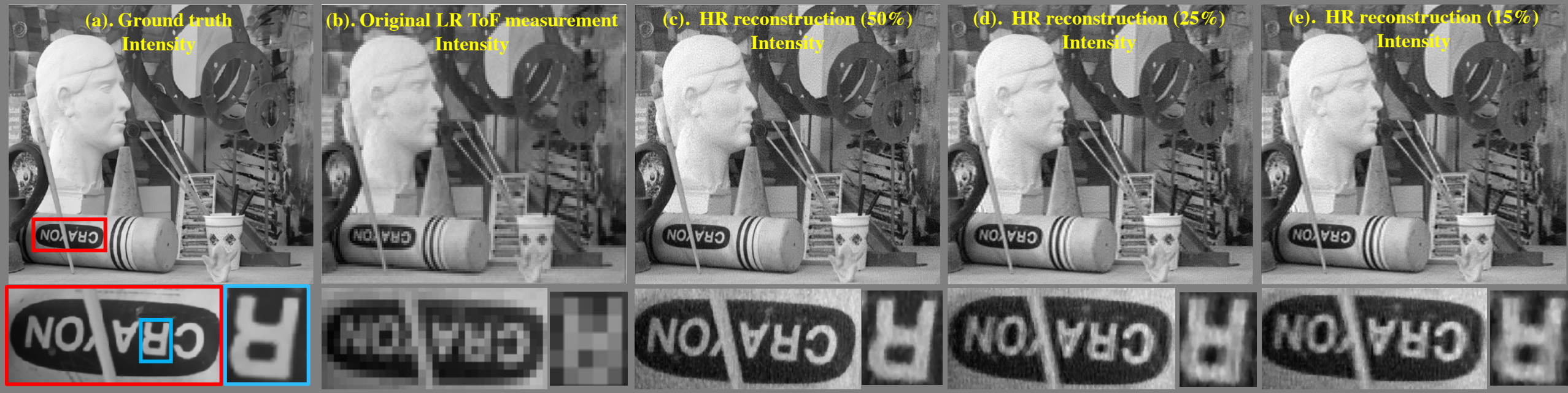
Optimization

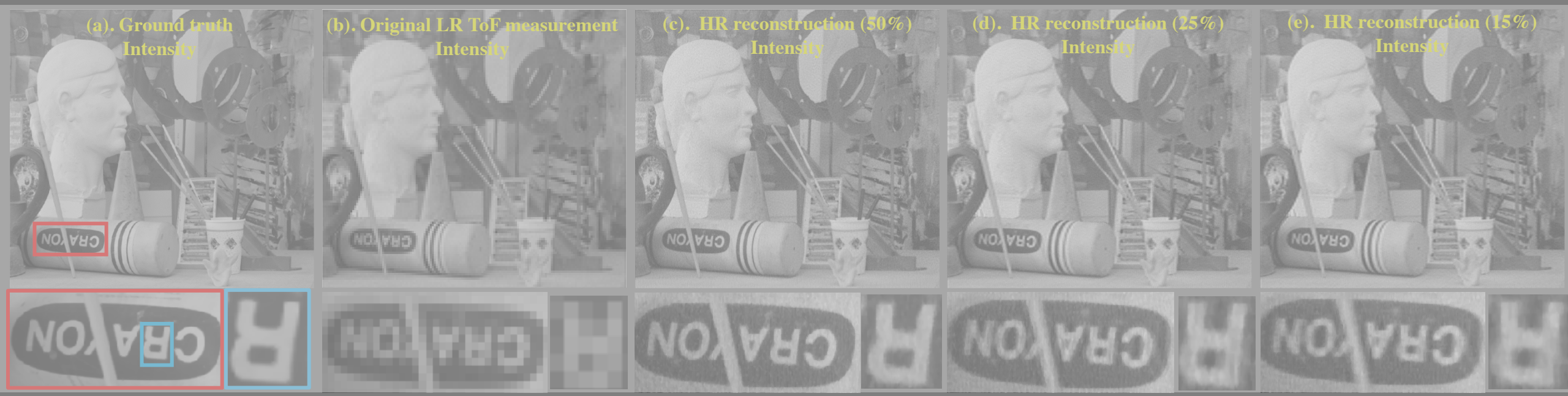
$$\hat{x} = \arg\min_{x} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda\Phi(x)$$

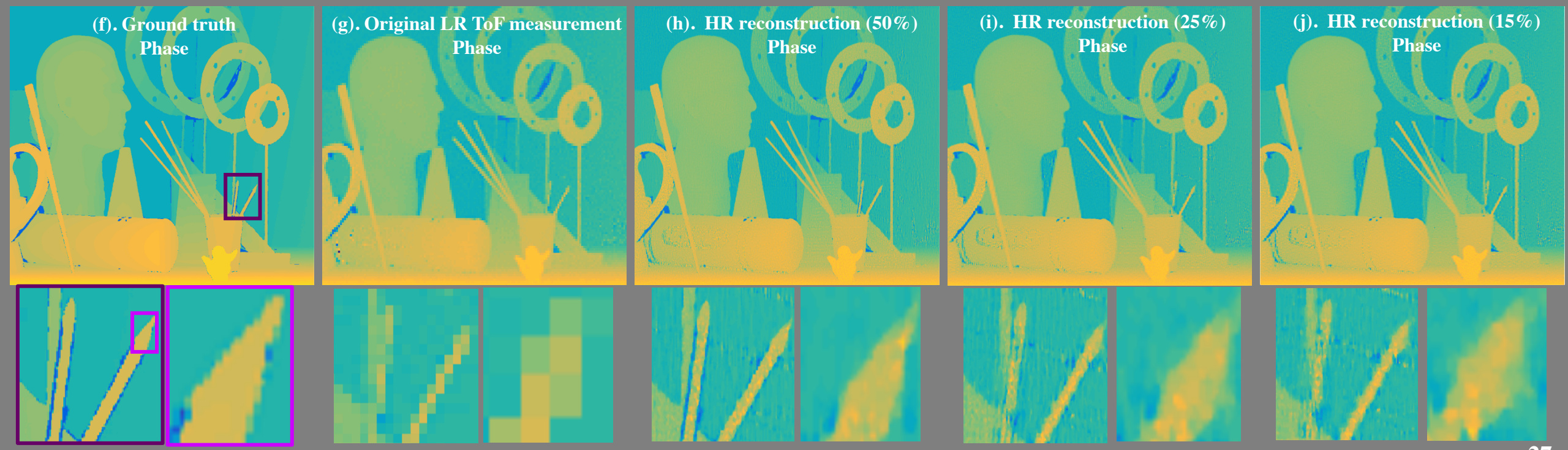
$$\Phi(\mathbf{x}) = \mathrm{TV}(\mathbf{x}) = \sum_{i} \sqrt{|G_u(x_i)|^2 + |G_v(x_i)|^2}$$

Simulation

- High resolution scene
  - Middlebury 3D Datasets
  - Size: 1140x912
- $\mathbf{A}_t$ – simulated system $\mathbf{A} = \mathbf{C}\mathbf{M}_t$
  - $\mathbf{M}_t$ – Hadamard multiplexing patterns on DMD at **t**-th measurement
  - **C** – mapping matrix, defined as spatial down-sampling by averaging
- Low resolution measurements
  - Simulated through via $\mathbf{y}_t = \mathbf{A}_t\mathbf{x}$
  - Size: 120×153


(a). Ground truth
Intensity
(b). Original LR ToF measurement
Intensity
(c). HR reconstruction (50%)
Intensity
(d). HR reconstruction (25%)
Intensity
(e). HR reconstruction (15%)
Intensity

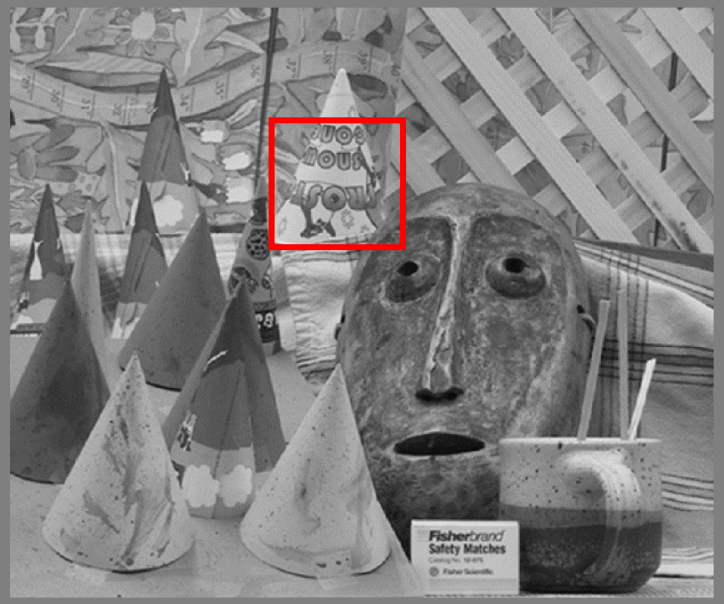

(a). Ground truth Intensity
(b). Original LR ToF measurement Intensity
(c). HR reconstruction (50%) Intensity
(d). HR reconstruction (25%) Intensity
(e). HR reconstruction (15%) Intensity

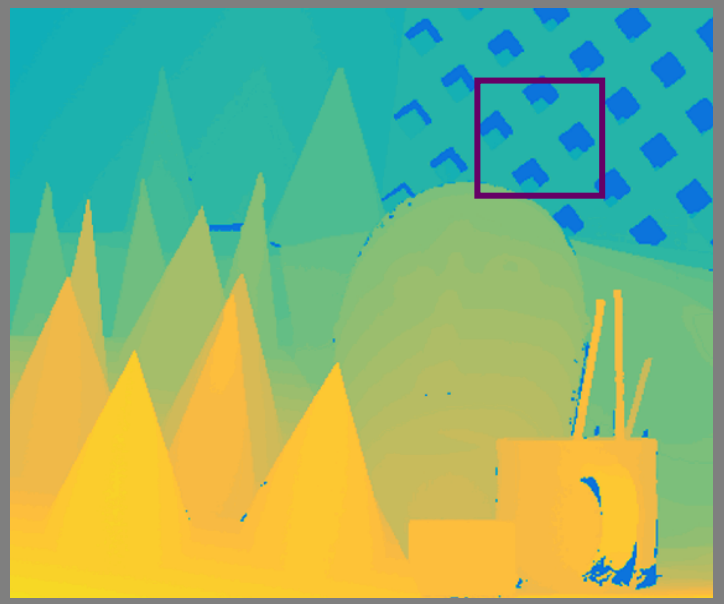

(f). Ground truth Phase
(g). Original LR ToF measurement Phase
(h). HR reconstruction (50%) Phase
(i). HR reconstruction (25%) Phase
(j). HR reconstruction (15%) Phase

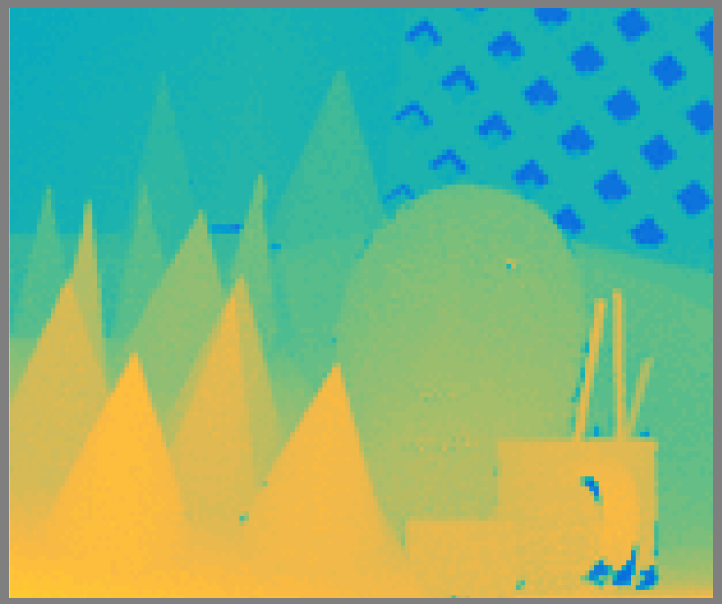
Ground truth

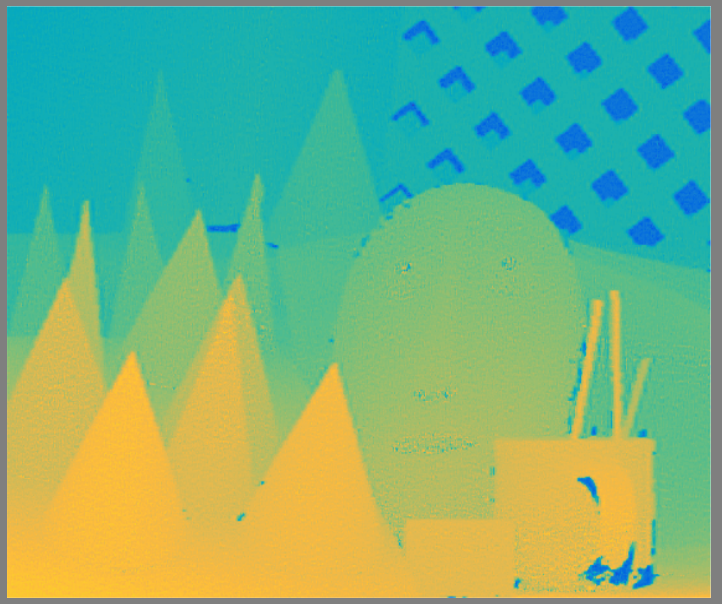
Original LR ToF measurement

HR reconstruction (15%)





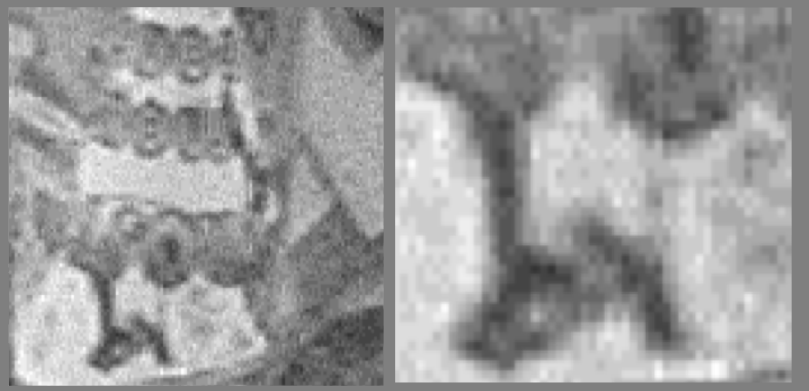


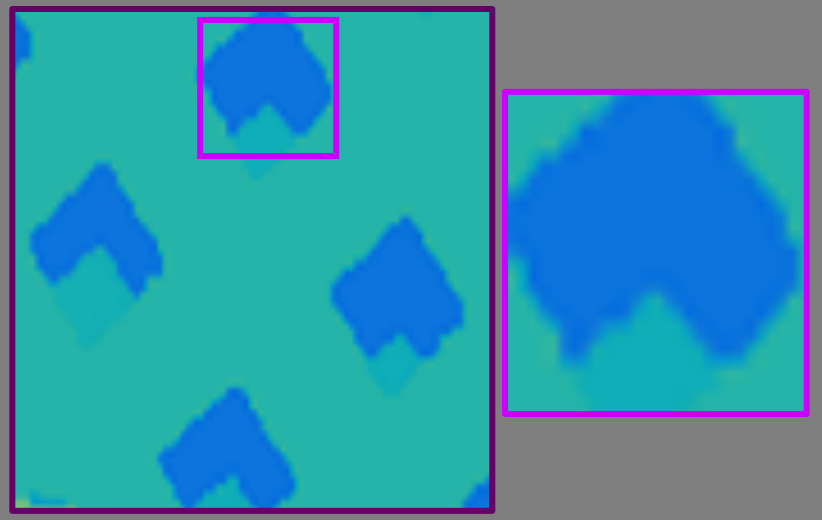


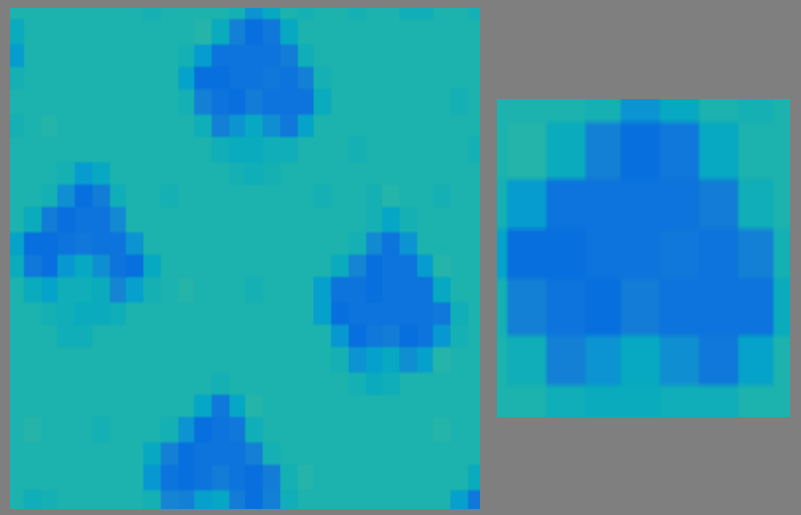
Ground truth

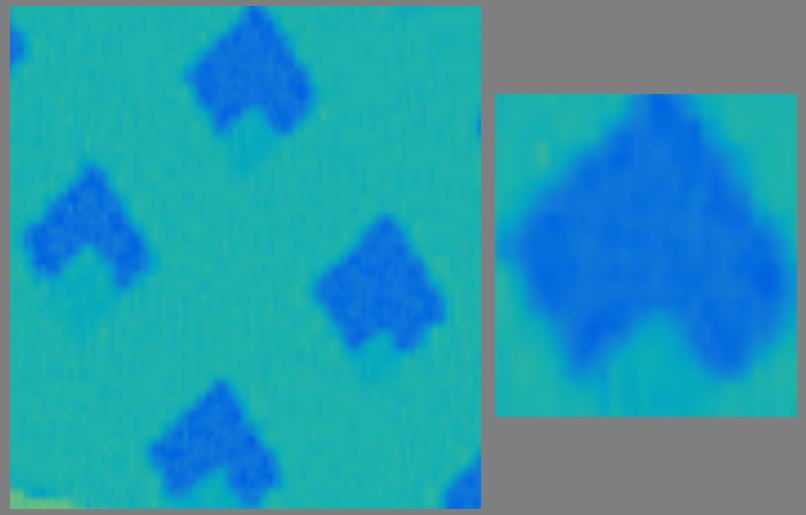
Original LR ToF measurement

HR reconstruction (15%)

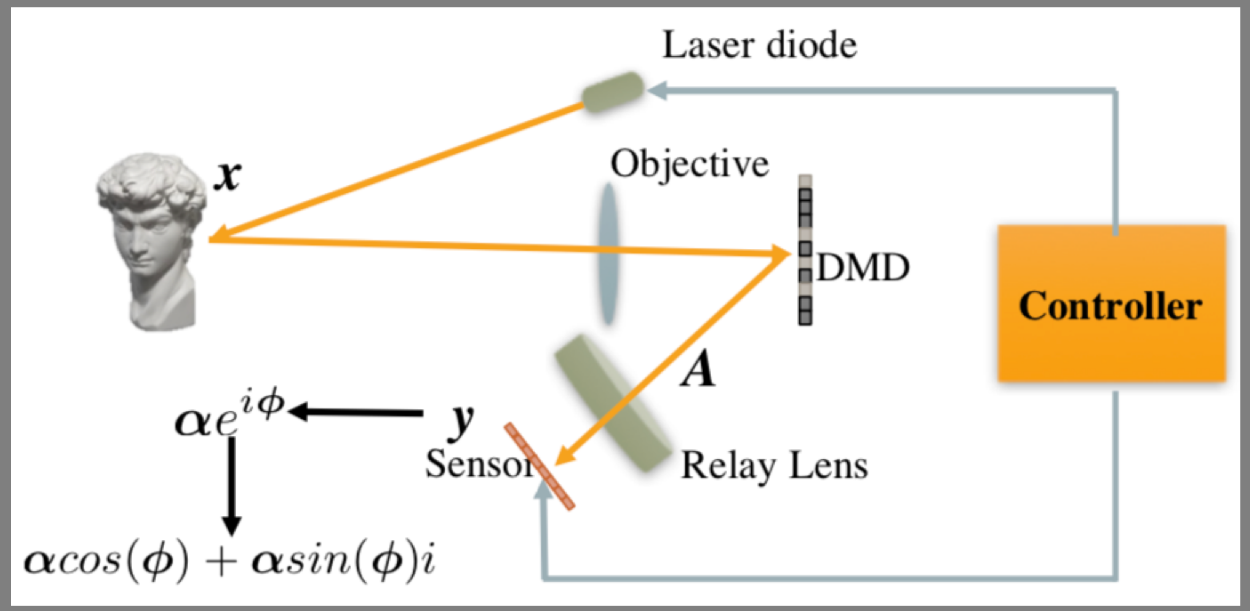


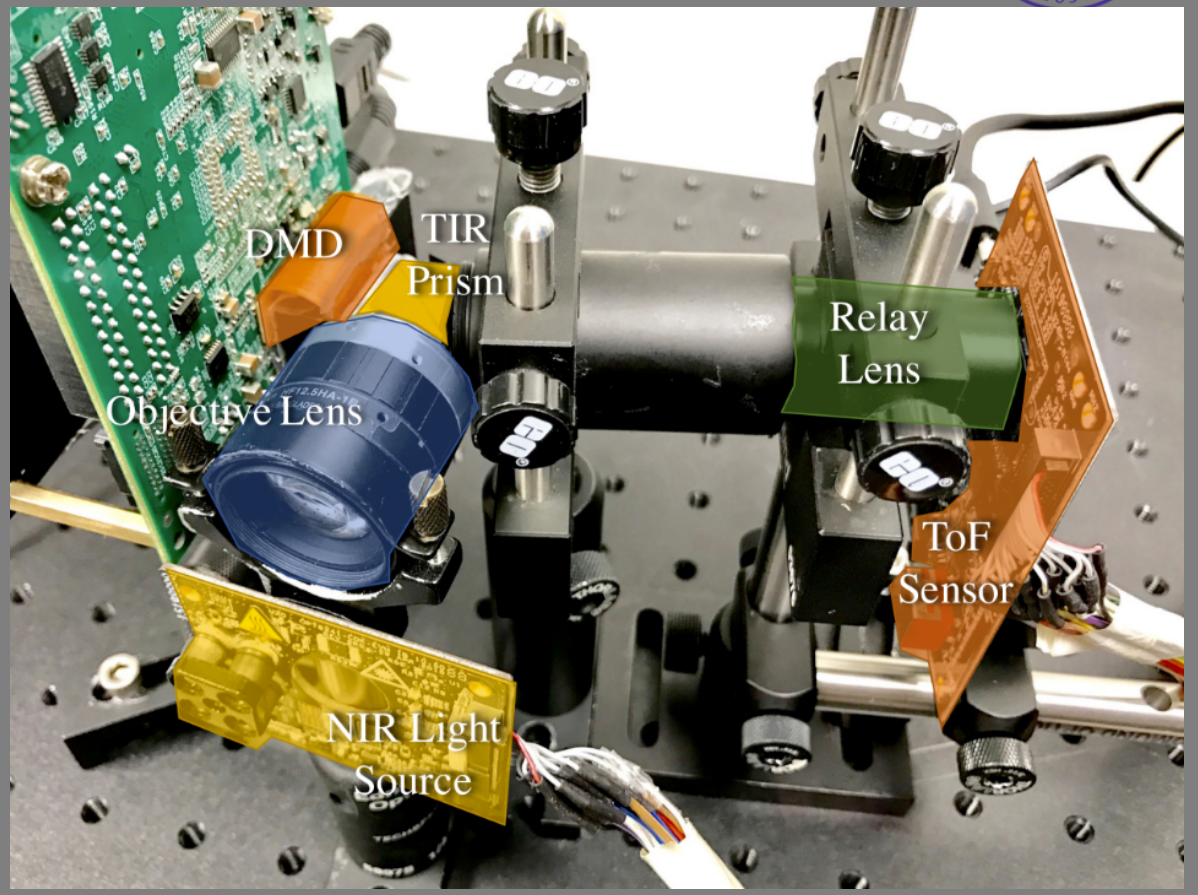






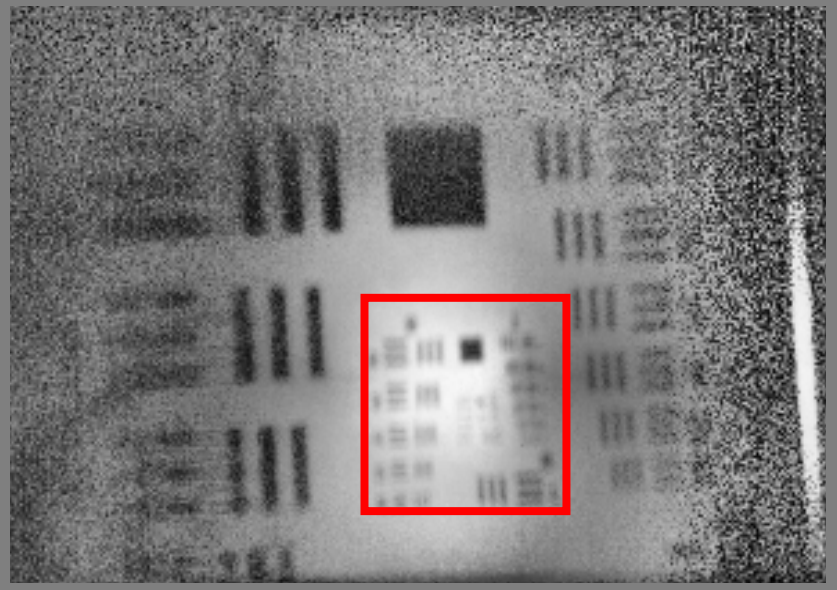
Prototype



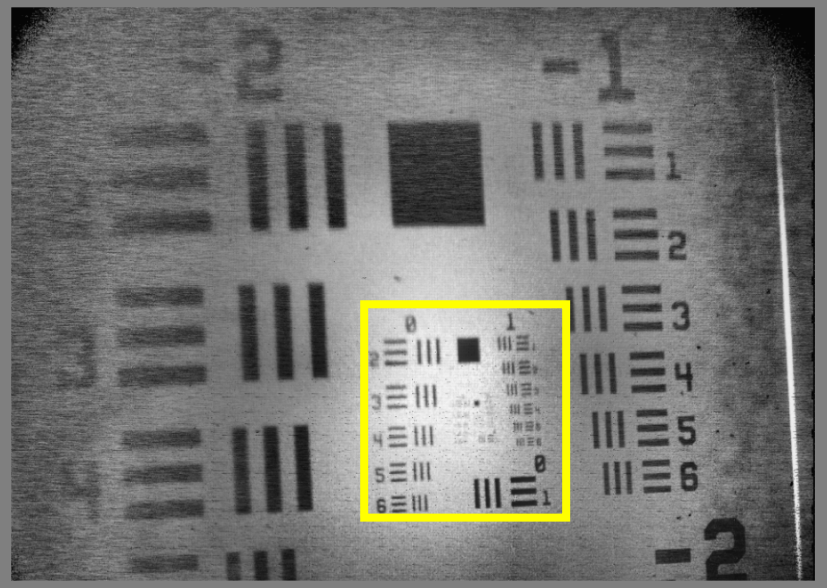


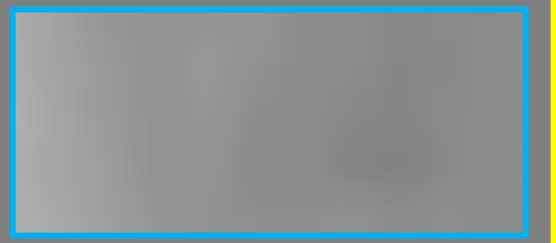
- Texas Instrument DLP 4500: 1140×912 pixels
- ToF camera: Texas Instrument OPT 8241 320×240 (186×200 pixels)

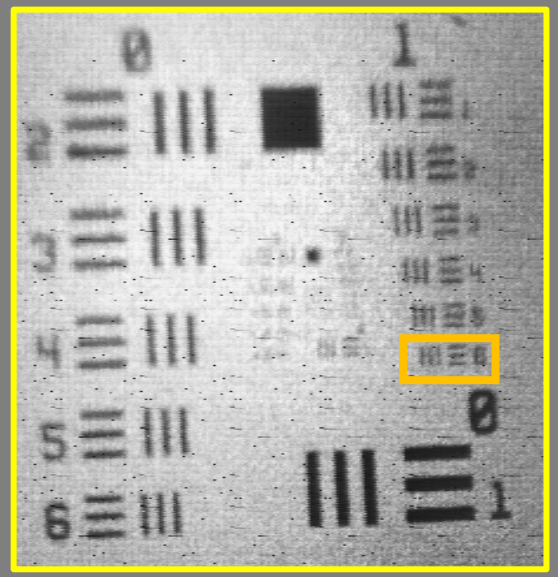
Original LR ToF Measurement



Pixel scanning

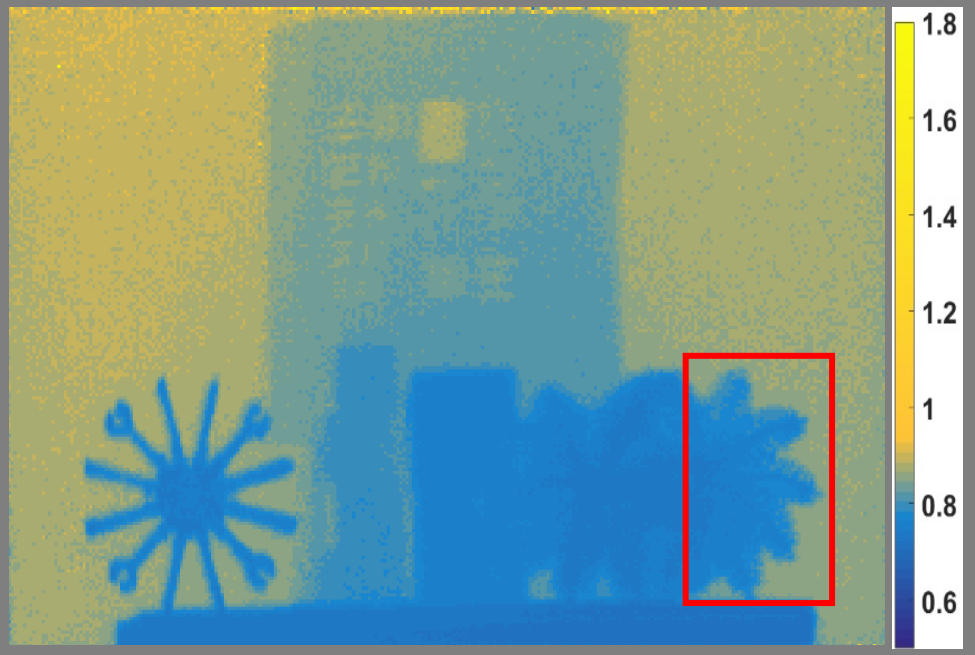




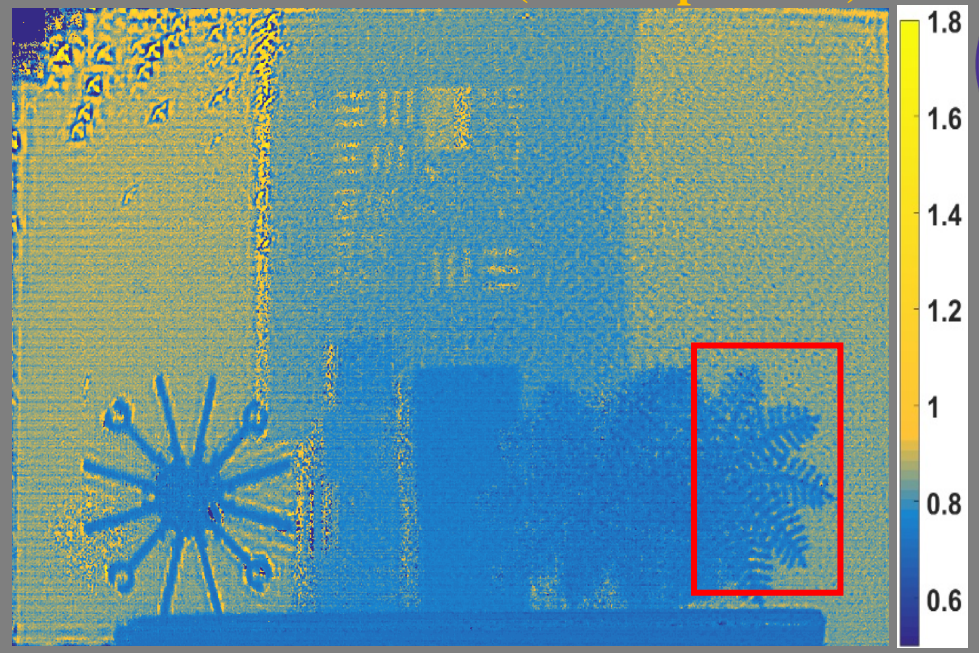


3D scene

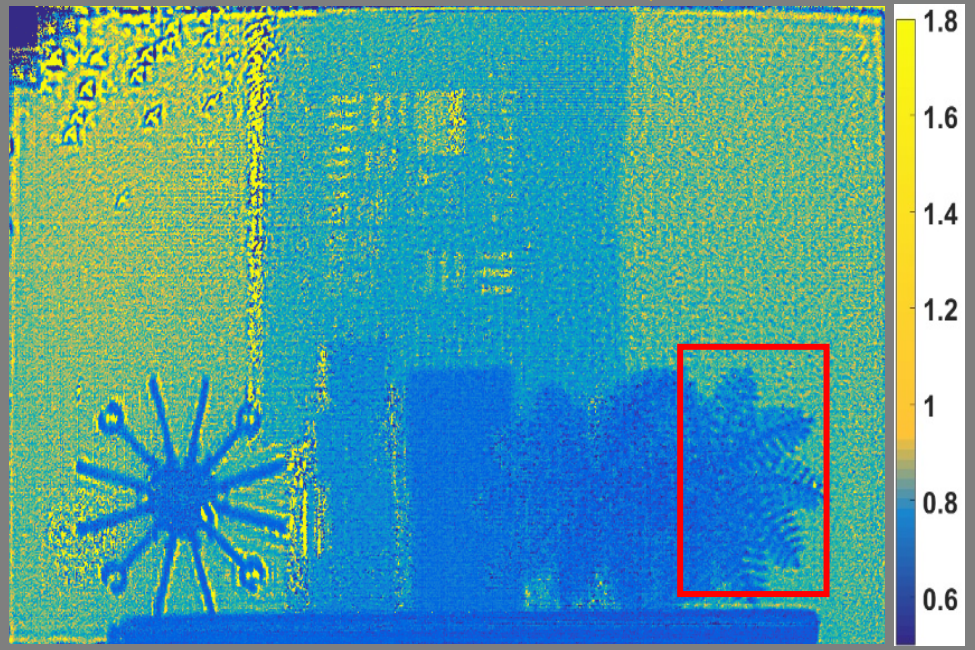
Original LR ToF measurement



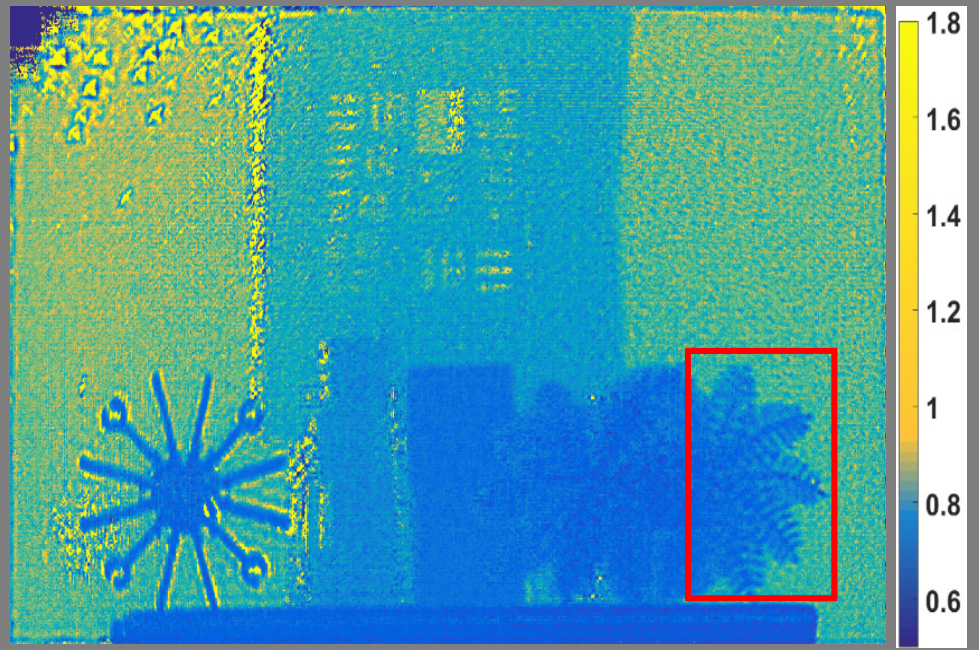
HR Reconstruction (no compression)

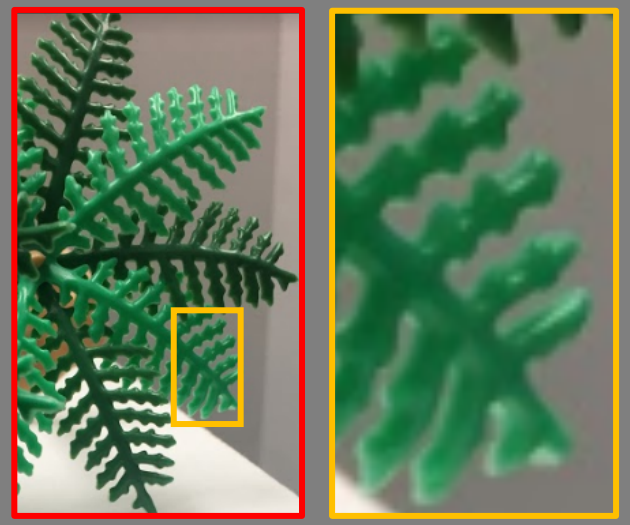


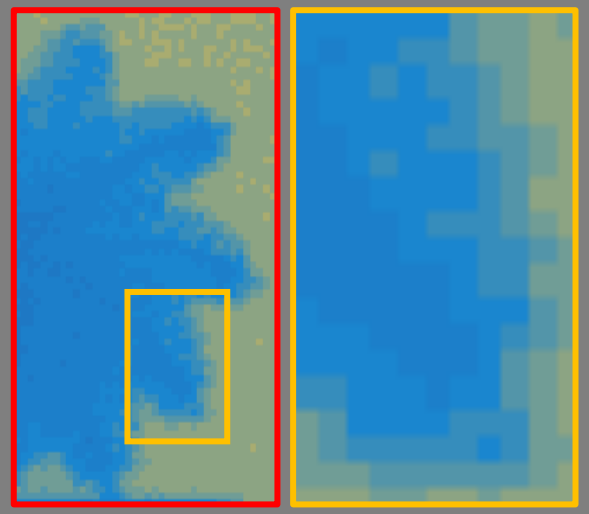
HR Reconstruction (0.6)

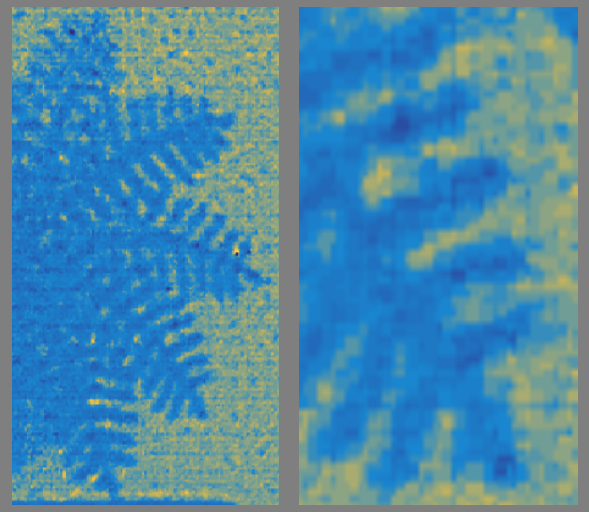


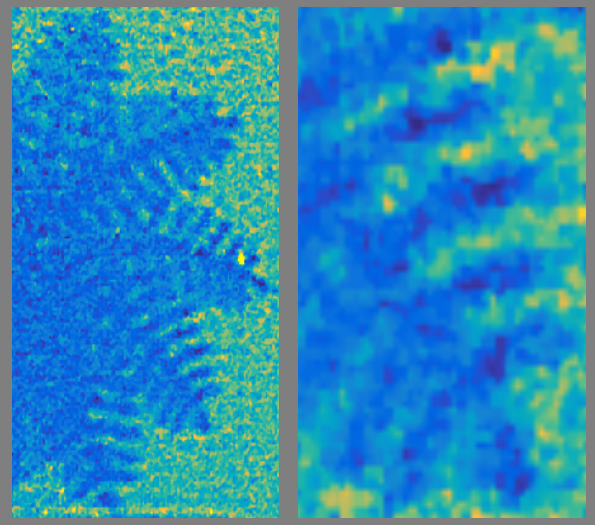
HR Reconstruction (0.25)

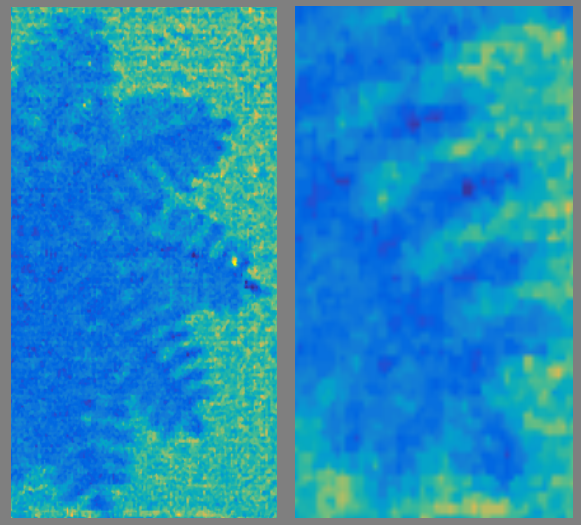










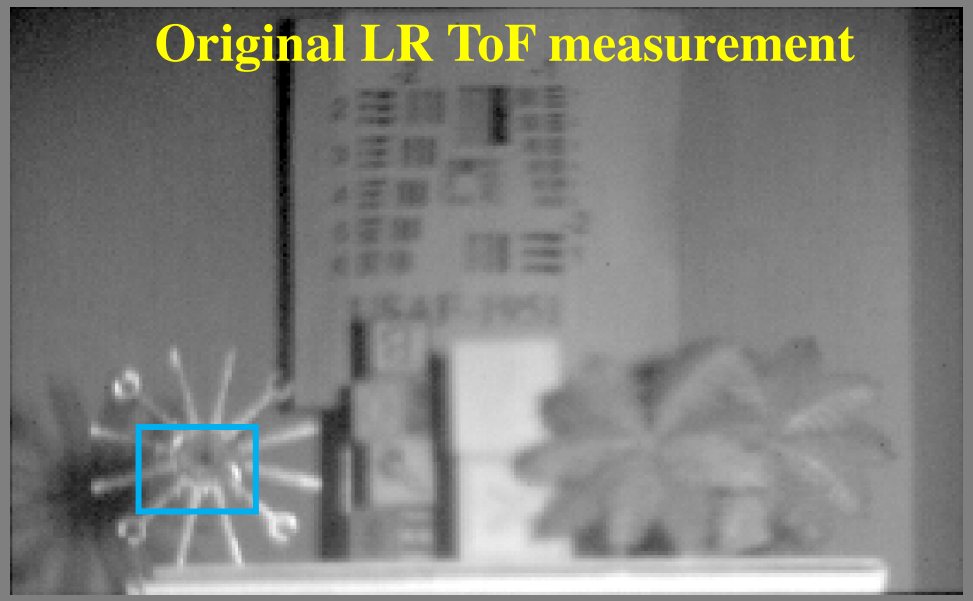


Native

No compression

CS: 0.6

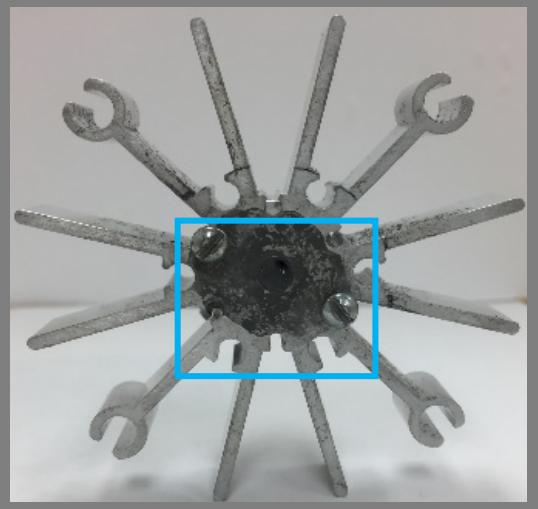
CS: 0.25

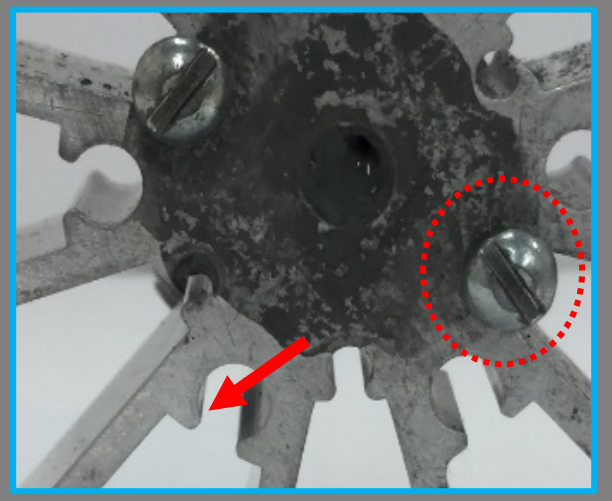

Original LR ToF measurement

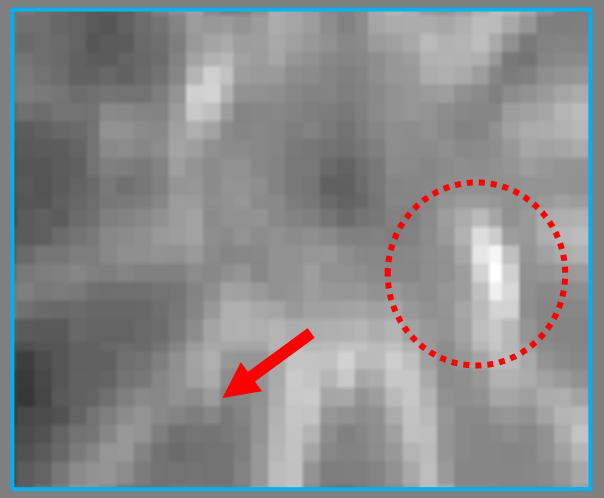

HR reconstruction (no compression)

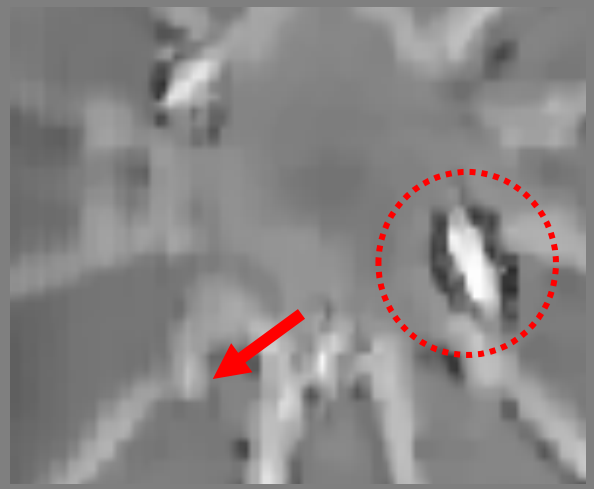

HR reconstruction (60%)

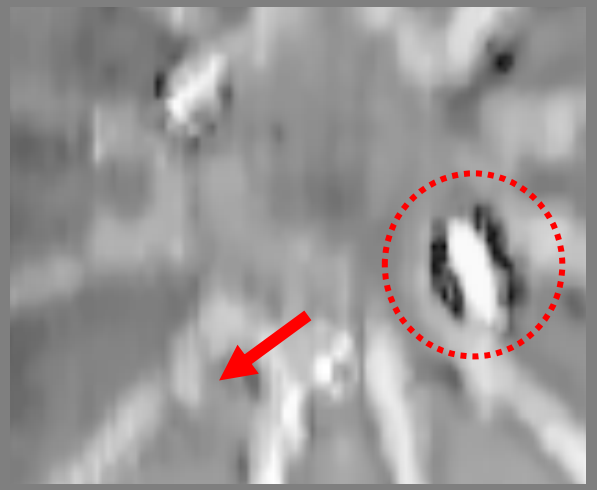

HR reconstruction (25%)











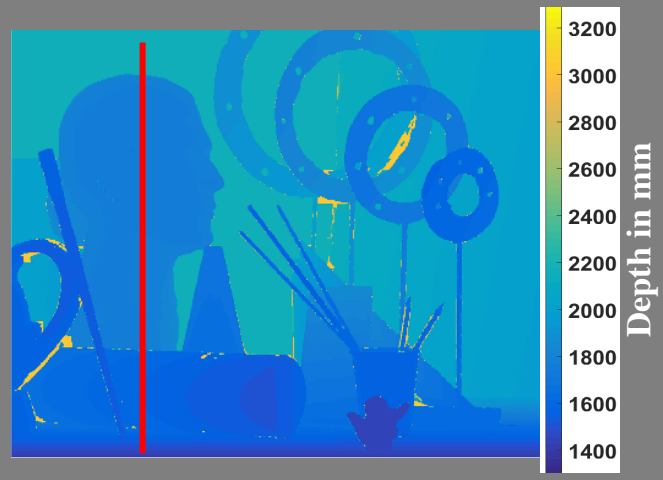
Native

No compression

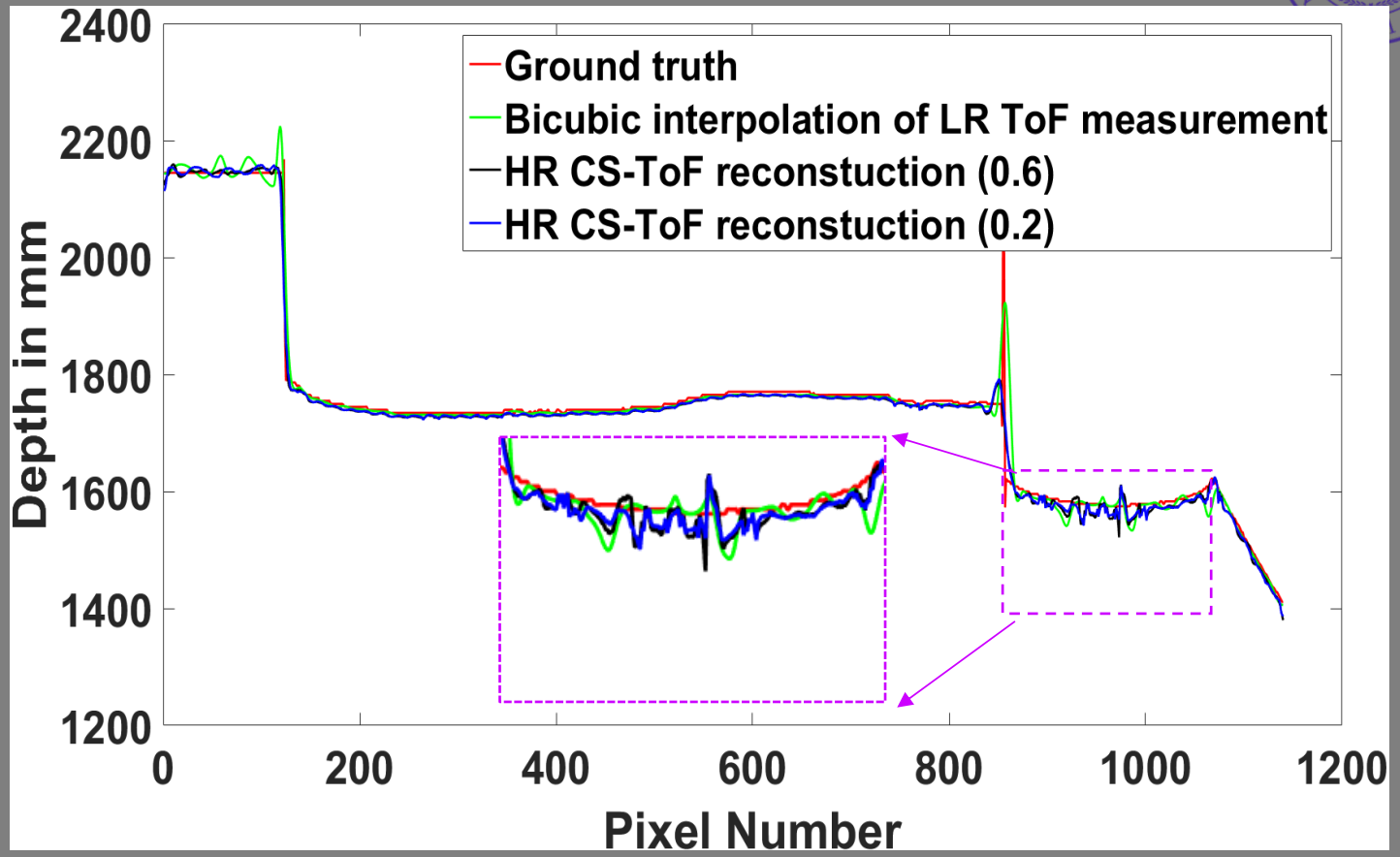
CS: 0.6

CS: 0.25

Depth resolution

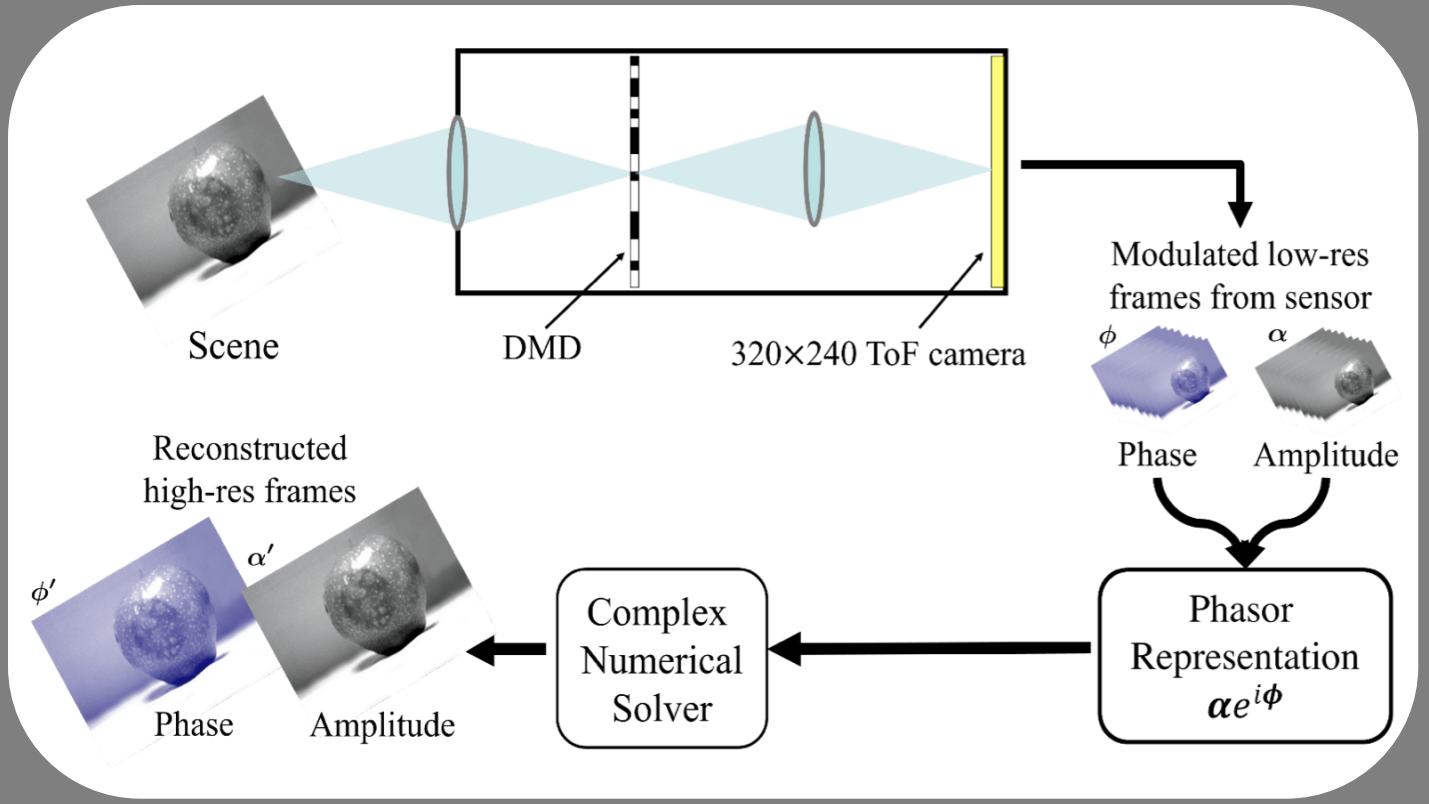


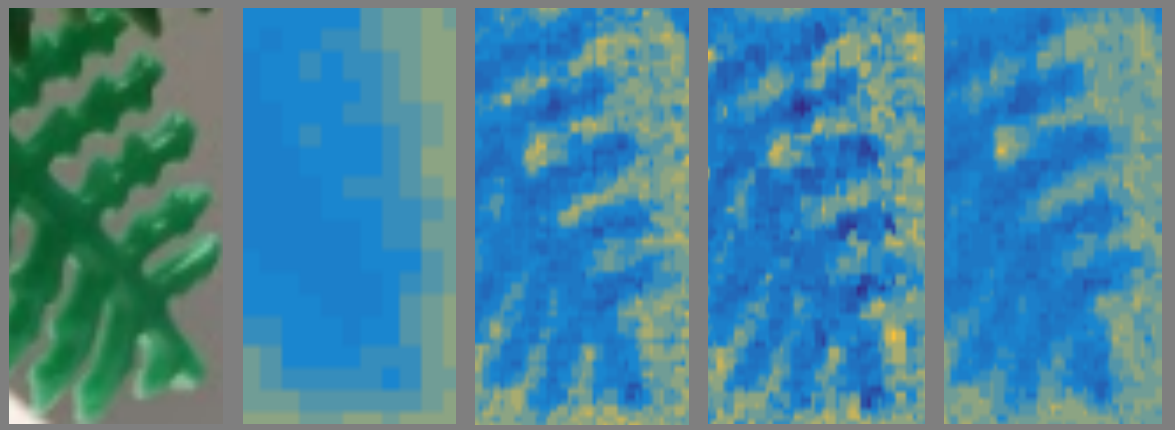
Ground truth depth



Depth values with 25dB SNR

- No depth resolution improvement
- Depth resolution is better than bicubic interpolation


Scene
DMD
320×240 ToF camera
Modulated low-res frames from sensor
ϕ
α
Phase
Amplitude
Phasor Representation
$\alpha e^{i\phi}$
Complex Numerical Solver
Reconstructed high-res frames
ϕ'
α'
Phase
Amplitude



Project Page

http://compphotolab.northwestern.edu/project/cs-tof-high-resolution-compressive-time-of-flight-imaging/