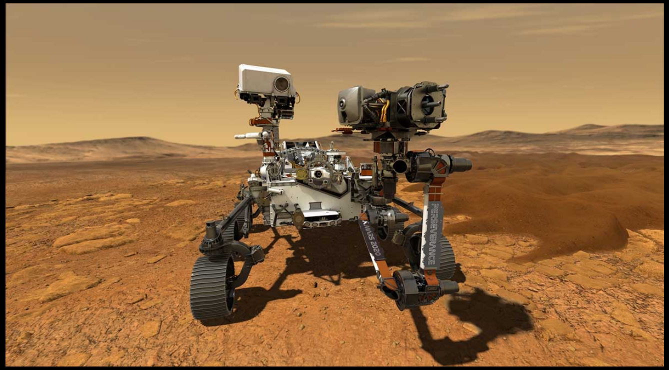
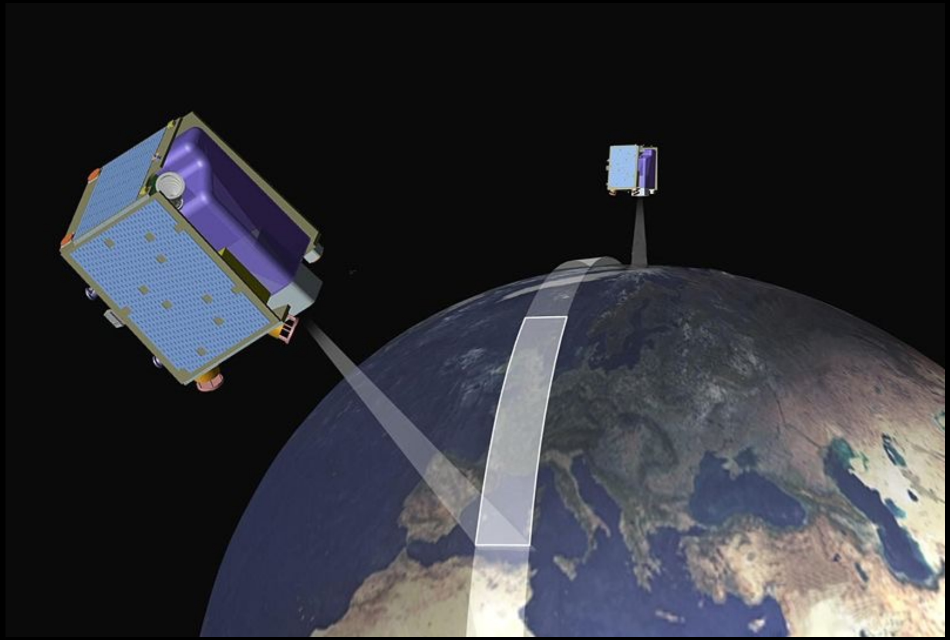
Physics-Based Visual Computing for Efficient 3D Vision and Sensing

David B. Lindell
University of Toronto


WAYMO
7ZGY053



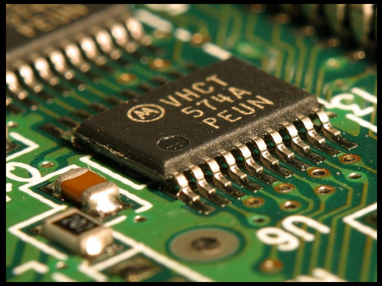


Physics-Based Visual Computing

Photon Interactions

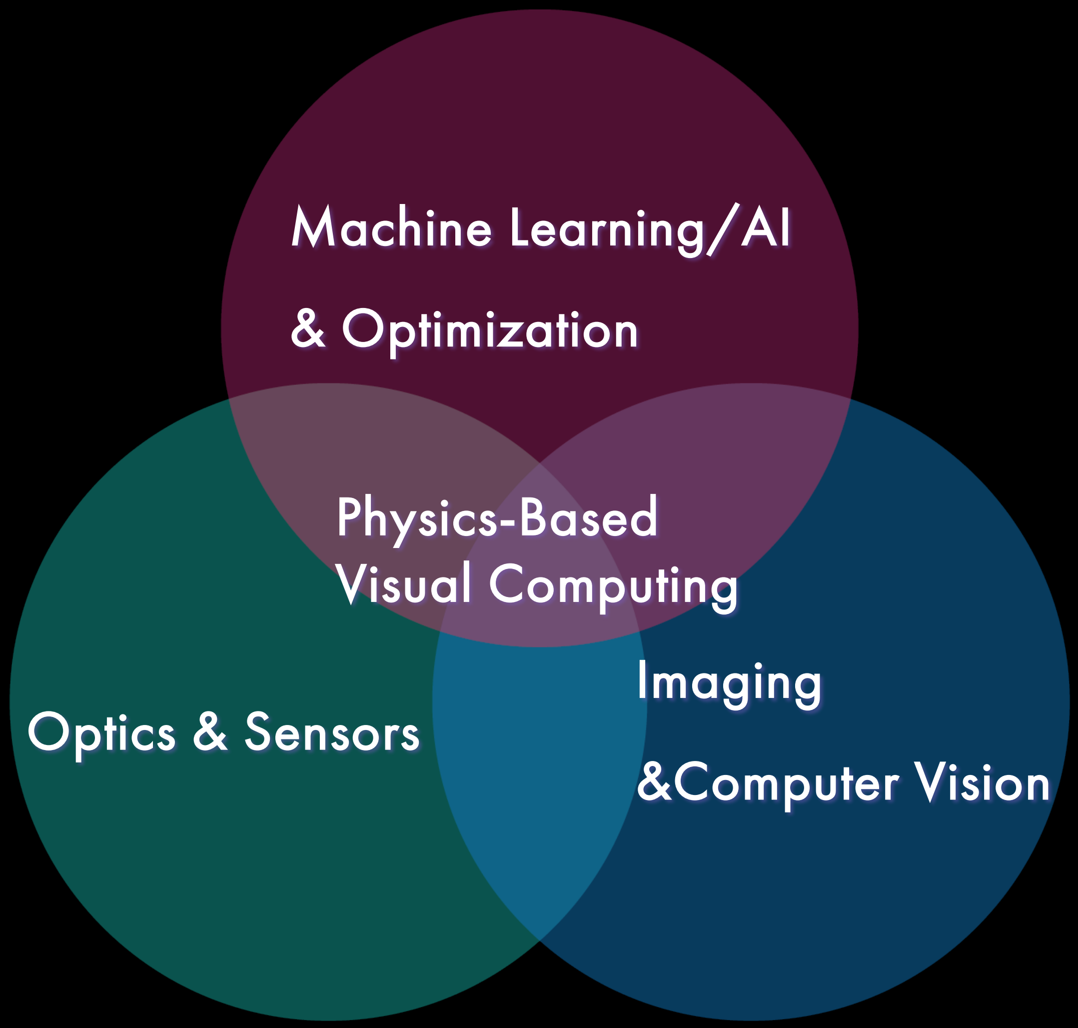
time of flight
polarization
spectrum
coherence
angle
spatial statistics
…

Optics, Sensors, Algorithms

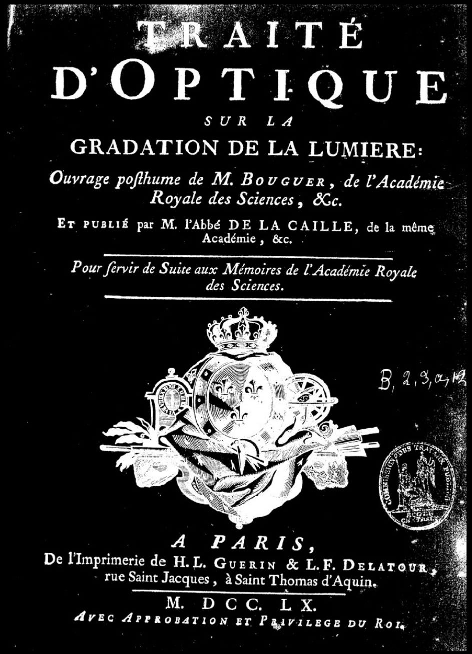


“Superhuman” Visual Computing

3D geometry, lighting, reflectance, material properties, motion, segmentation, semantics, behavior, etc.


Machine Learning/AI
& Optimization
Physics-Based
Visual Computing
Optics & Sensors
Imaging
&Computer Vision

Pierre Bouguer



TRAITÉ
D'OPTIQUE
SUR LA
GRADATION DE LA LUMIERE:

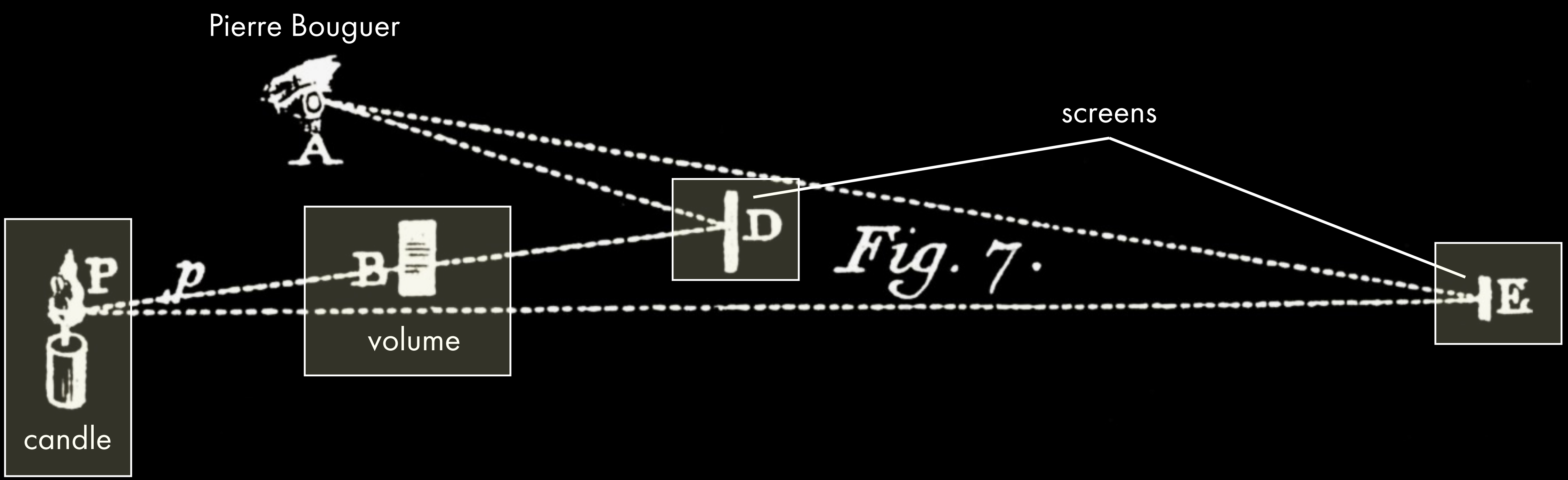
Ouvrage posthume de M. Bouguer, de l'Académie Royale des Sciences, &c.

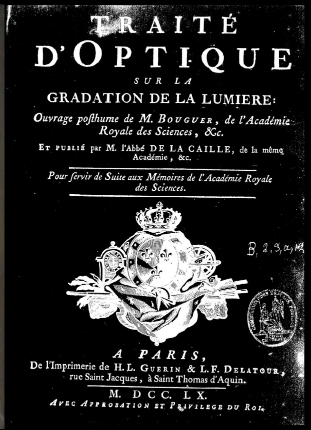
Et publié par M. l'Abbé DE LA CAILLE, de la même Académie, &c.

Pour servir de Suite aux Mémoires de l'Académie Royale des Sciences.

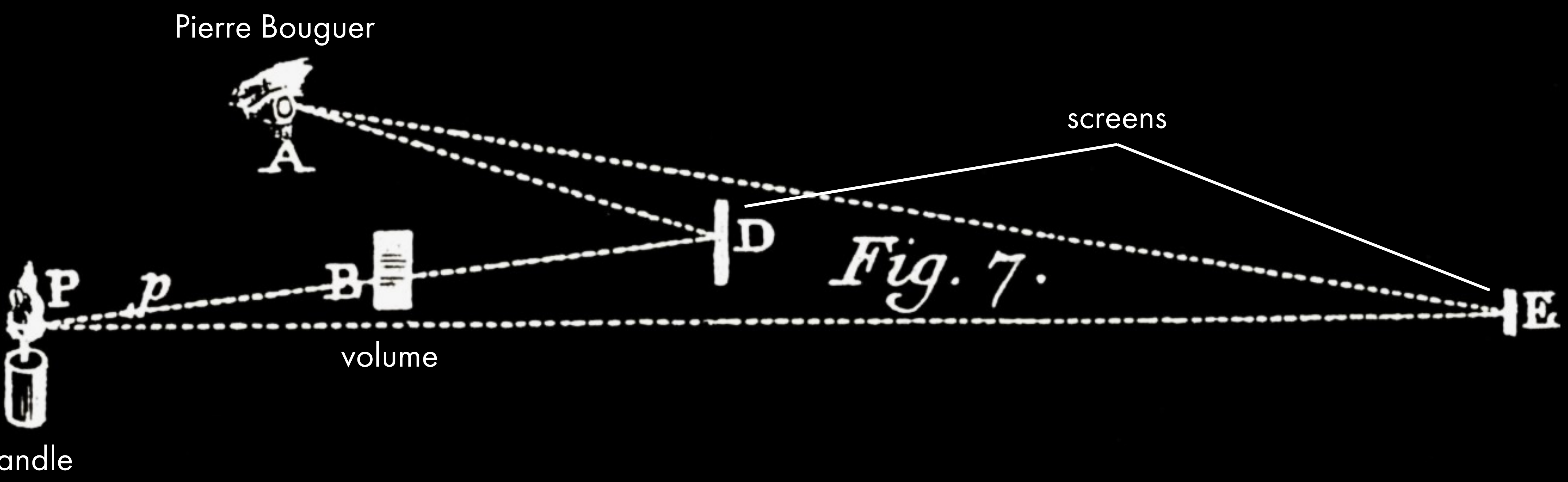
A PARIS,
De l'Imprimerie de H.L. Guerin & L.F. Delatour, rue Saint Jacques, à Saint Thomas d'Aquin.
M. DCC. LX.
Avec Approbation et Privilege du Roi.

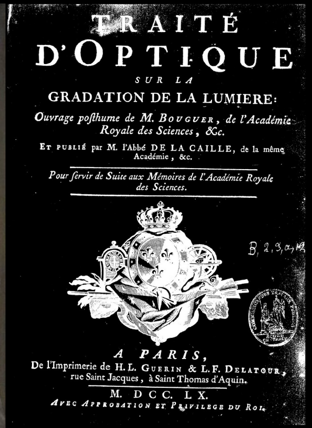
Treatise on Optical Gradations (1760)




TRAITÉ
D'OPTIQUE
SUR LA
GRADATION DE LA LUMIERE:
Ouvrage posthume de M. Bouguer, de l'Académie Royale des Sciences, &c.
Et publié par M. l'Abbé DE LA CAILLE, de la même Académie, &c.
Pour servir de Suite aux Mémoires de l'Académie Royale des Sciences.
A PARIS,
De l'Imprimerie de H.L. Guerin & L.F. Delatour, rue Saint Jacques, à Saint Thomas d'Aquin.
M. DCC. LX.
Avec Approbation et Privilege du Roi.

Treatise on Optical Gradations (1760)





TRAITÉ
D'OPTIQUE
SUR LA
GRADATION DE LA LUMIERE:

Ouvrage posthume de M. Bouguer, de l'Académie Royale des Sciences, &c.

ET PUBLIÉ par M. l'Abbé DE LA CAILLE, de la même Académie, &c.

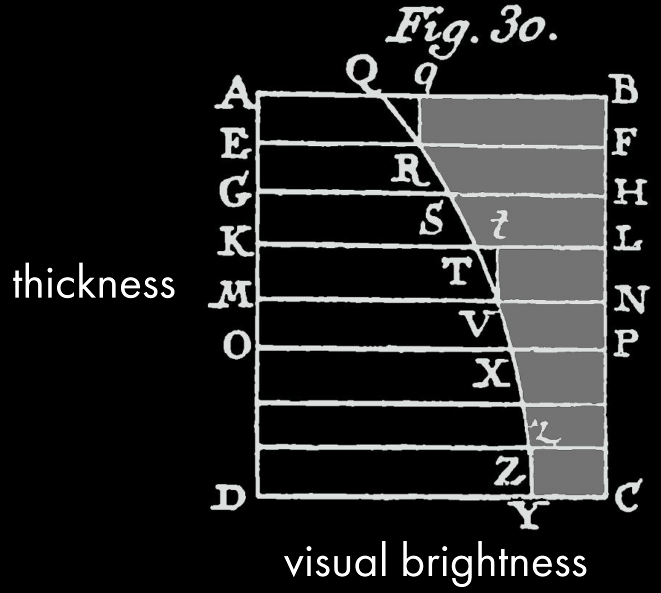
Pour servir de Suite aux Mémoires de l'Académie Royale des Sciences.

A PARIS,
De l'Imprimerie de H.L. Guerin & L.F. Delatour, rue Saint Jacques, à Saint Thomas d'Aquin.

M. DCC. LX.

Avec Approbation et Privilege du Roi.

Treatise on Optical Gradations (1760)



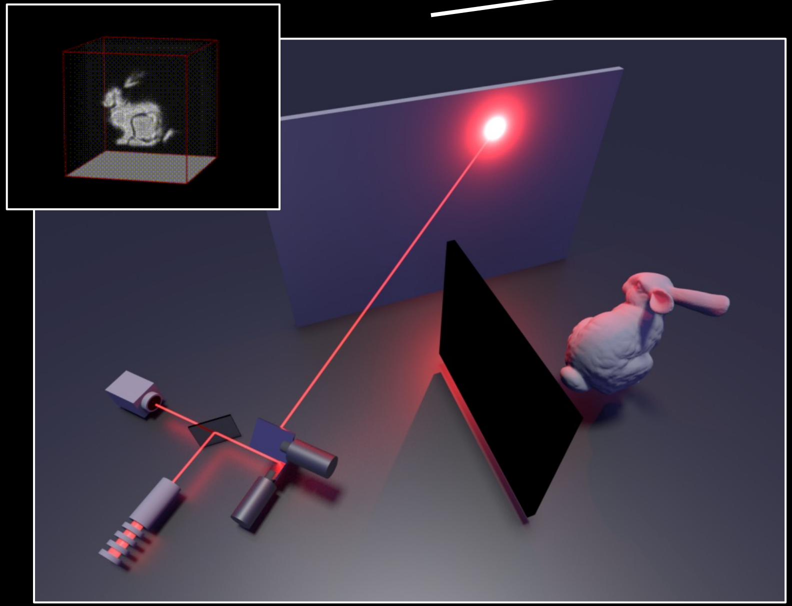
Radiative Transfer

$$\underbrace{(\omega \cdot \nabla) L(\mathbf{x}, \omega)}_{\text{radiance in direction } \omega} = \underbrace{-\sigma_t(\mathbf{x}) L(\mathbf{x}, \omega)}_{\text{scattering/absorption}} + \underbrace{L_e(\mathbf{x}, \omega)}_{\text{emission}} + \underbrace{\sigma_s(\mathbf{x}) \int_{\mathcal{S}^2} f_p(\mathbf{x}, \omega, \omega') L(\mathbf{x}, \omega') \mathrm{d}\omega'}_{\text{in-scattering}}$$

Overview

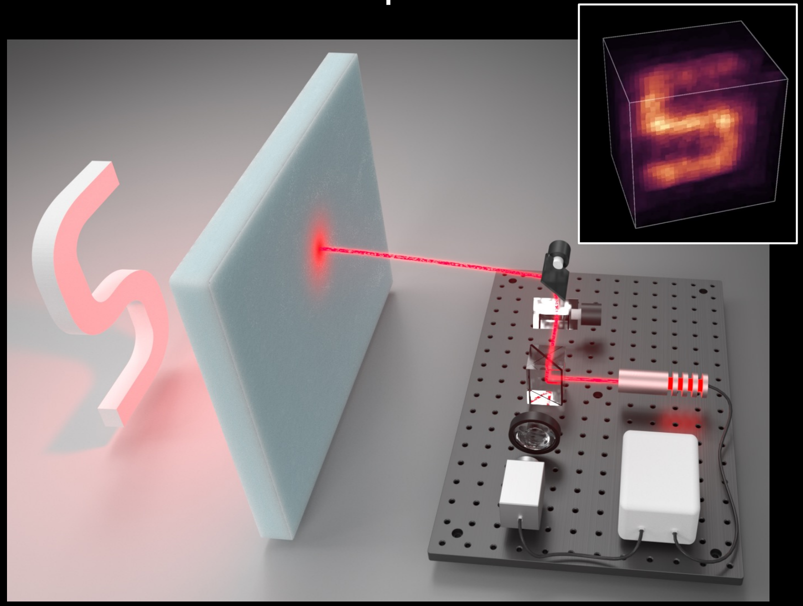
Radiative Transfer

$$(\omega \cdot \nabla)L(\mathbf{x}, \omega) = -\sigma_t(\mathbf{x})L(\mathbf{x}, \omega) + L_e(\mathbf{x}, \omega) + \sigma_s(\mathbf{x}) \int_{S^2} f_p(\mathbf{x}, \omega, \omega')L(\mathbf{x}, \omega')\mathrm{d}\omega$$



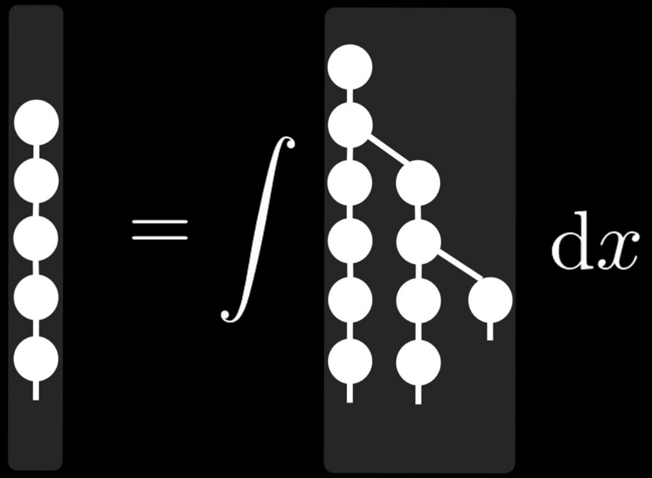
Non-Line-of-Sight Imaging

Nature '18
SIGGRAPH '19
CVPR '19
ACM Trans. Graph. '20
CVPR '20
IEEE TCI '21



Imaging through Scattering Media

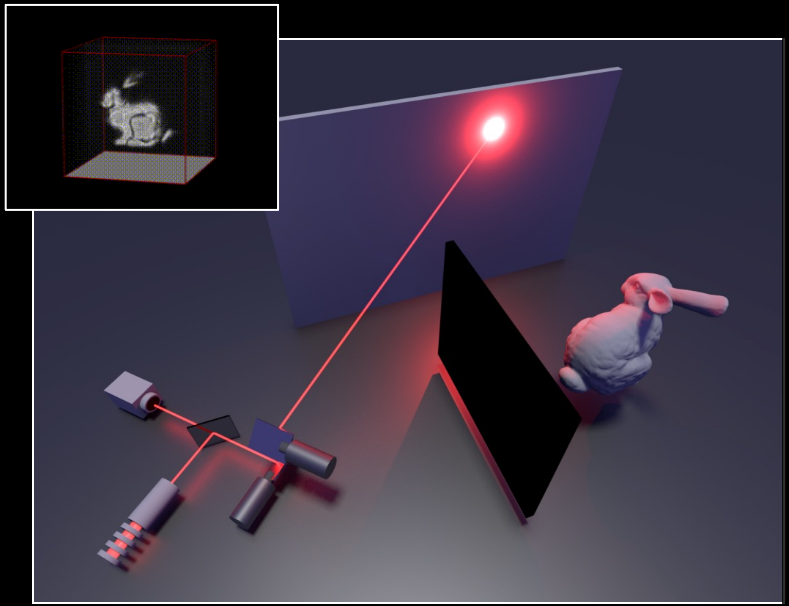
Nature Communications '20



Physics-based AI & Neural Rendering

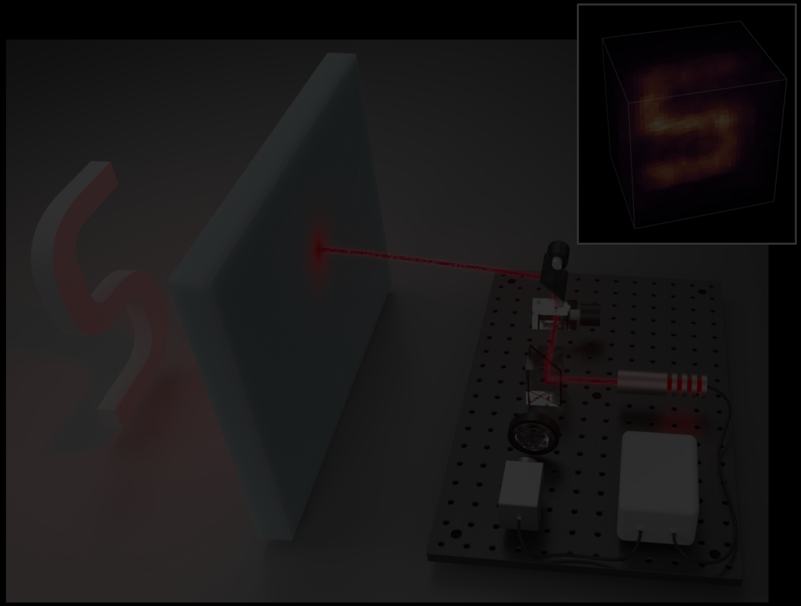
NeurIPS '20
CVPR '21
CVPR '22

Overview



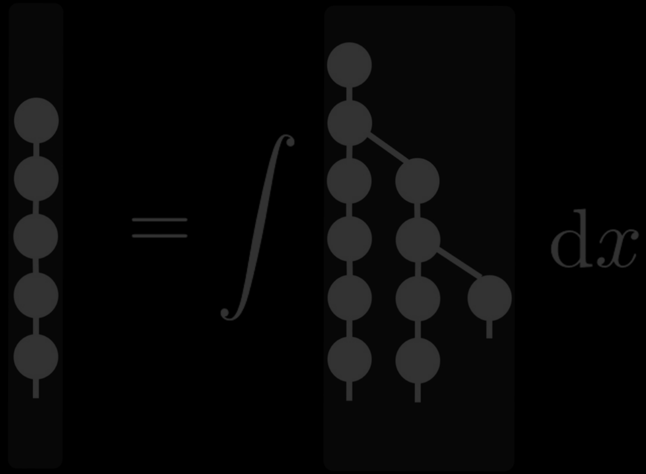
Non-Line-of-Sight Imaging

Nature '18
SIGGRAPH '19
CVPR '19
ACM Trans. Graph. '20
CVPR '20
IEEE TCI '21



Imaging through Scattering Media

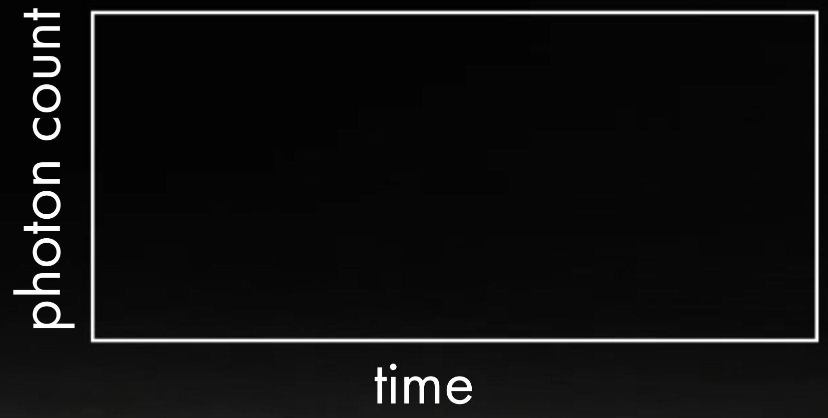
Nature Communications '20



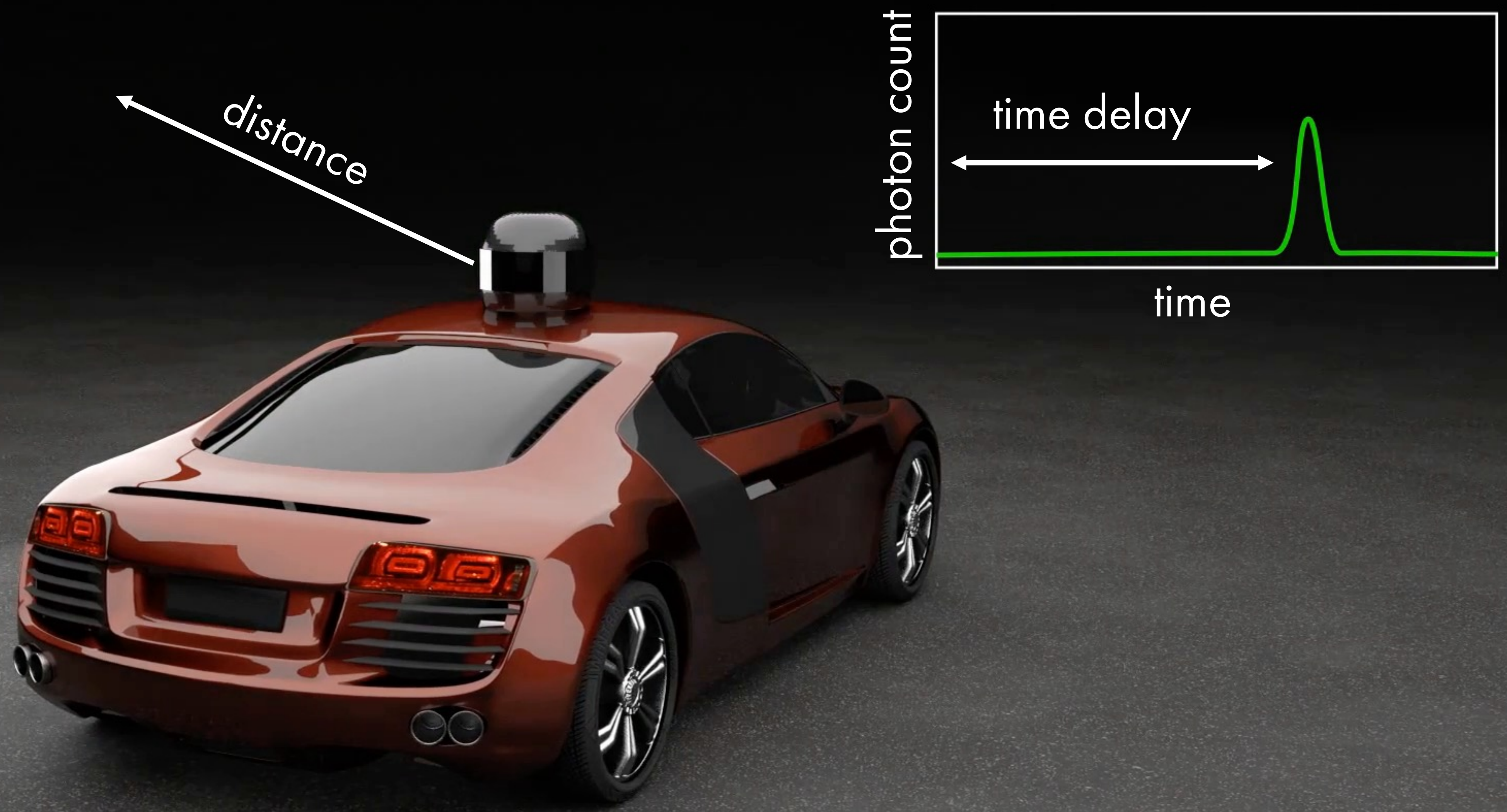
Physics-based AI & Neural Rendering

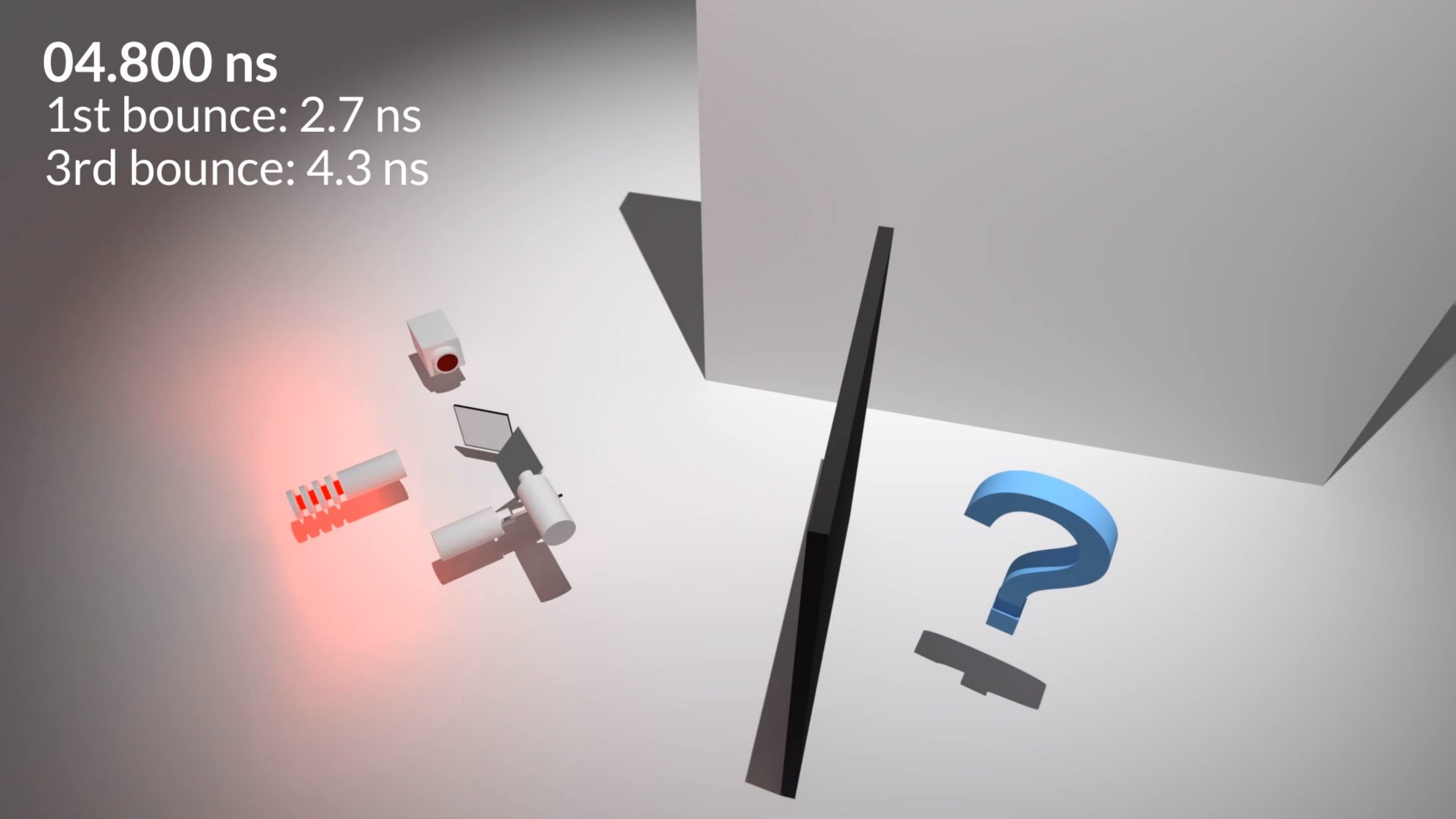
NeurIPS '20
CVPR '21
In submission '21

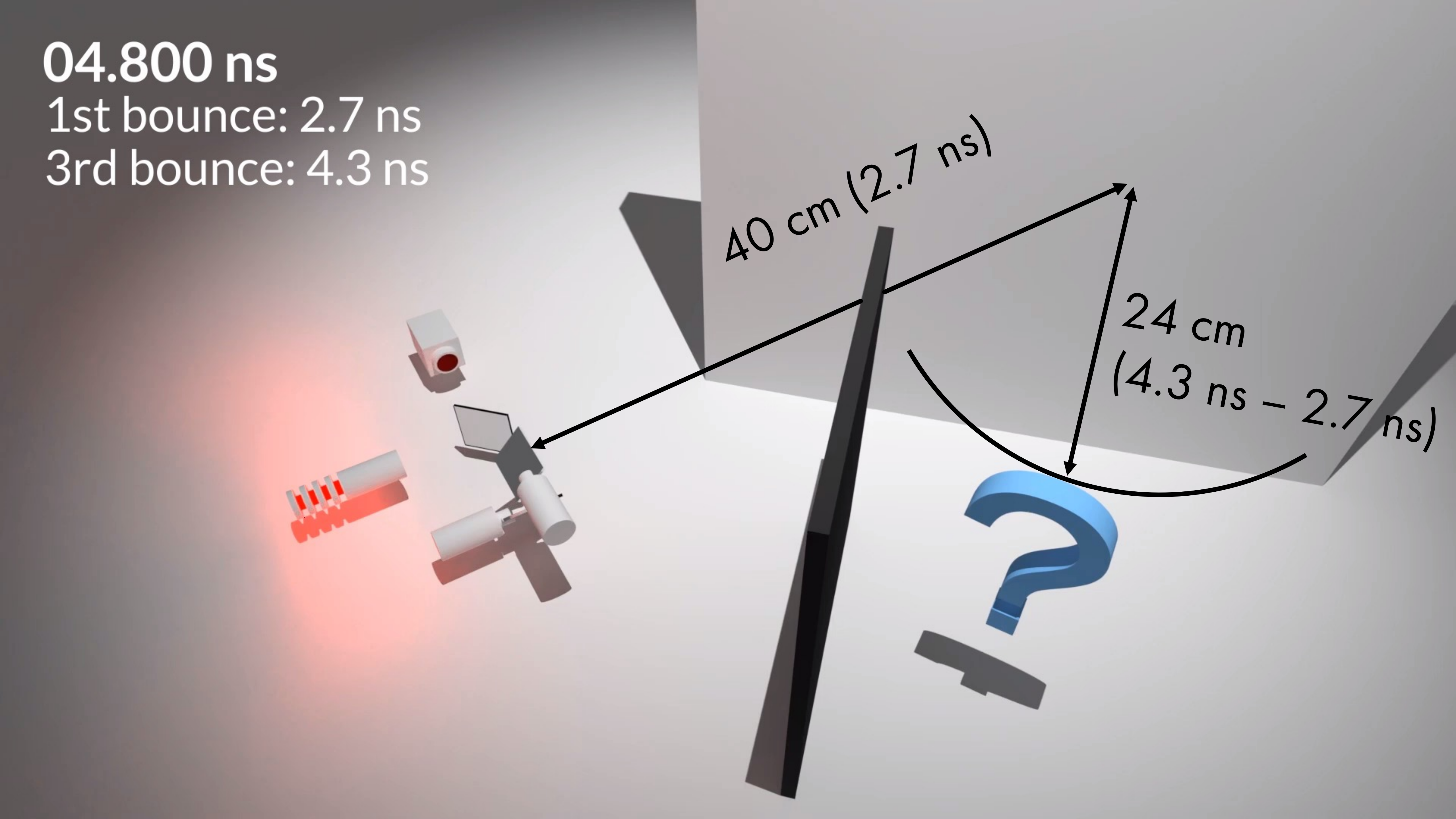
LIDAR



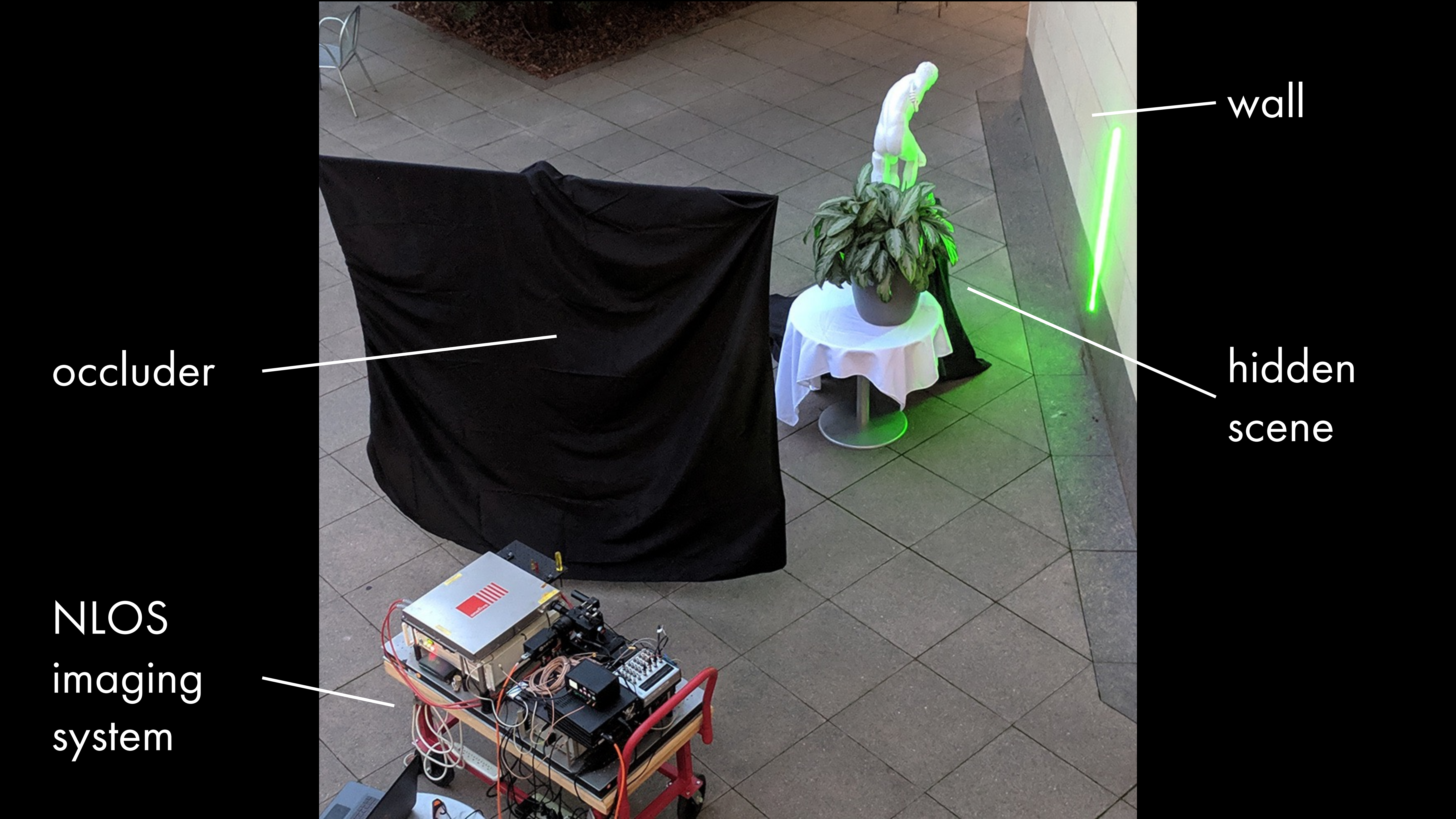
LIDAR



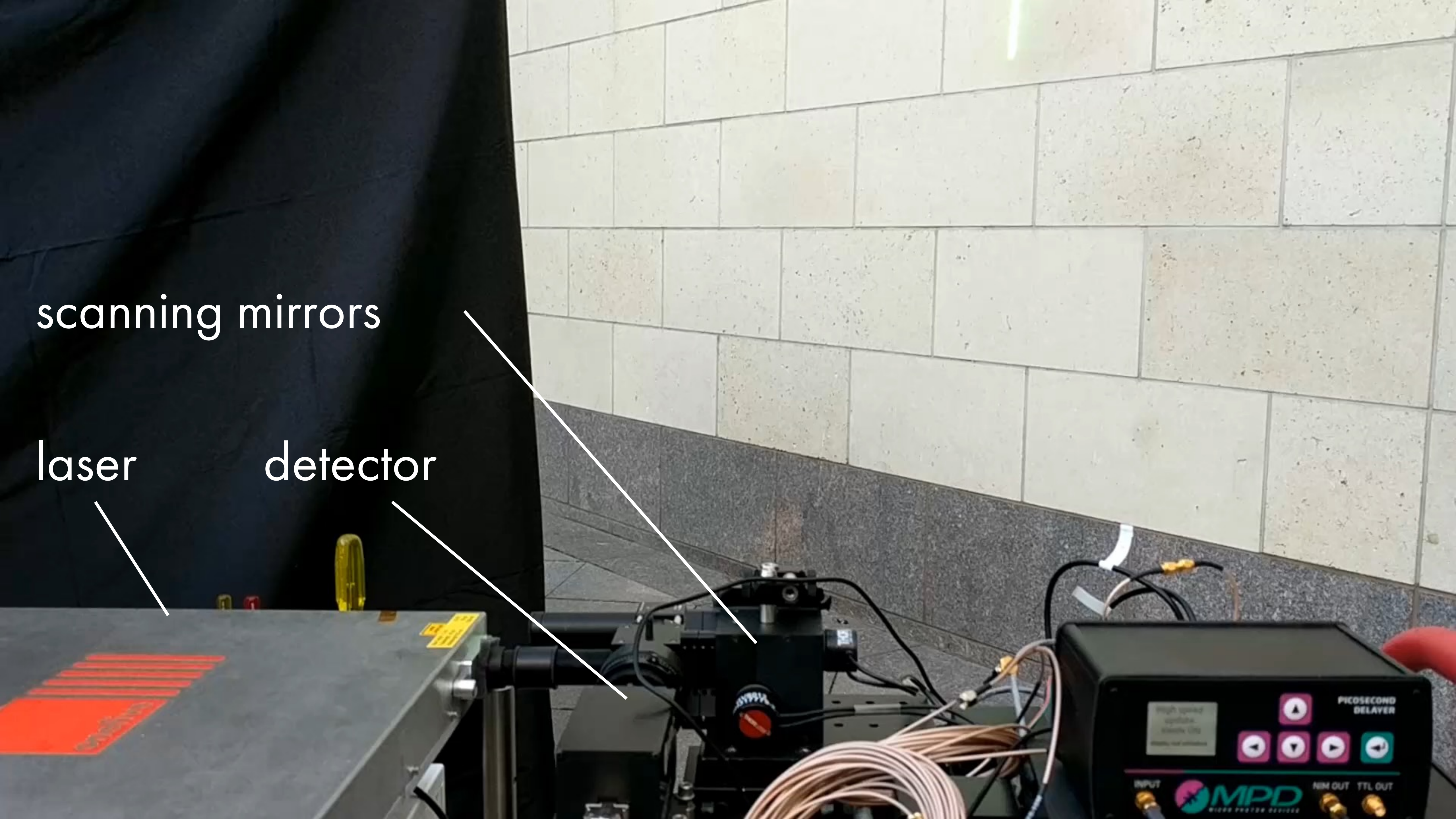


04.800 ns
1st bounce: 2.7 ns
3rd bounce: 4.3 ns



04.800 ns
1st bounce: 2.7 ns
3rd bounce: 4.3 ns
40 cm (2.7 ns)
24 cm
(4.3 ns – 2.7 ns)


wall
occluder
hidden scene
NLOS imaging system

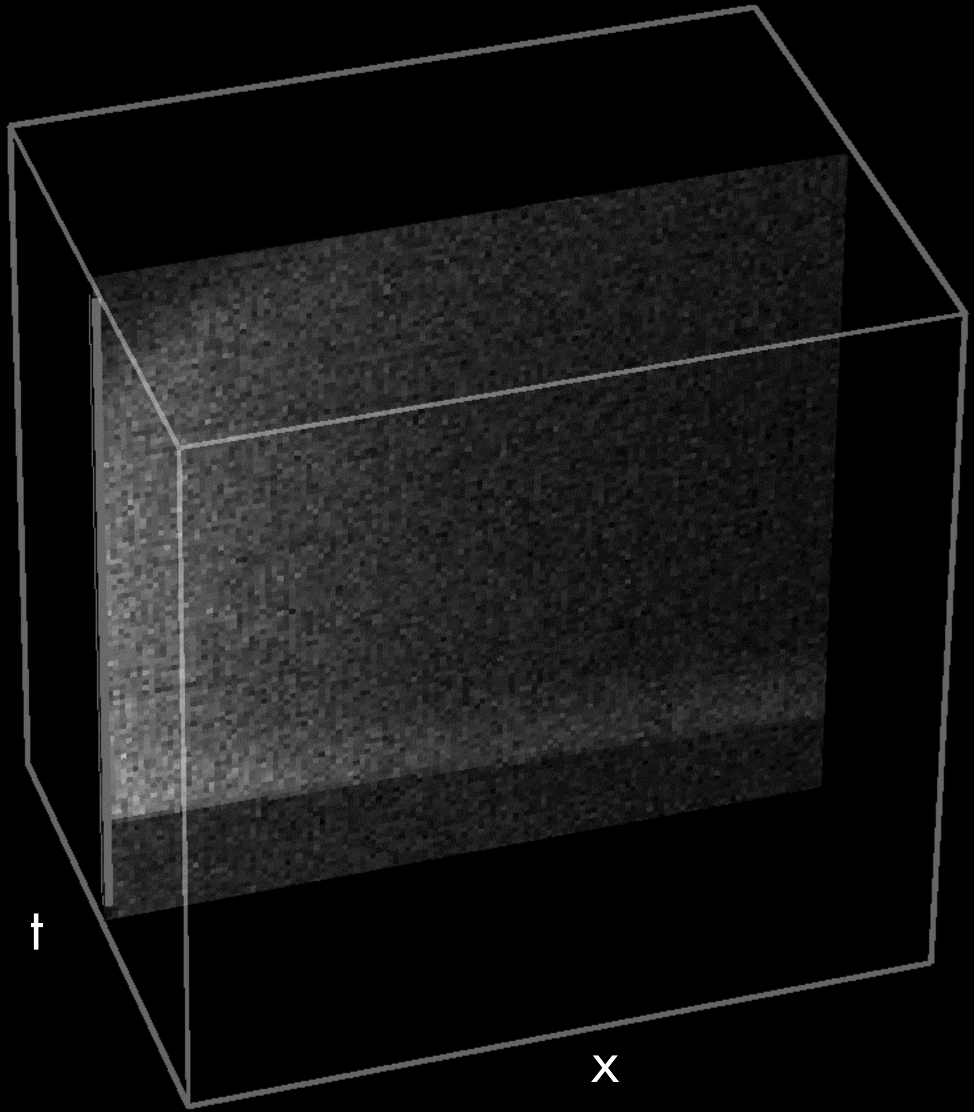

wall
occluder
hidden
scene
NLOS
imaging
system
Gated SPAD
Beam Splitter
Laser
Galvo Mirrors
Beam Dump


scanning mirrors
laser
detector
PICOSECOND DELAYER
INPUT
MPD
NIM OUT
TTL OUT


resolution: 128 x 128
area: 2 m × 2 m

scene photo

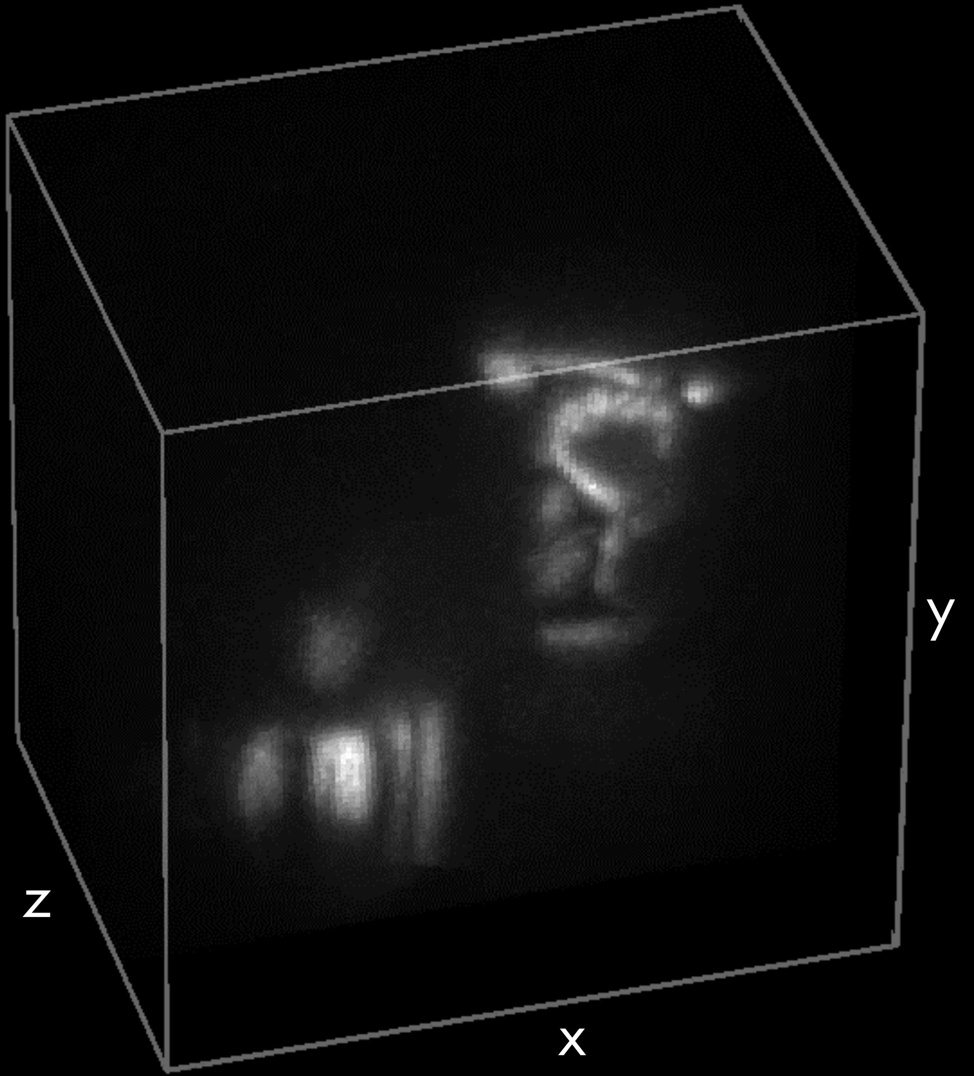
measurements



Lindell et al., SIGGRAPH 2019

scene photo

reconstruction



Dimensions: 2 m x 2 m x 1.5 m

Lindell et al., SIGGRAPH 2019

Challenges

1. Light efficiency / photon sensitivity

- weak signal from multiply scattered light
- emit as much light as possible - *fundamentally limited by eye safety* (in most applications)
- detect as much light as possible, ideally individual photons

Challenges

2. High-speed time stamping (determines accuracy)

- speed of light is ~300,000,000 m/s
- 1 m = 3.3 ns; 1 cm = 33 ps; 1 mm = 3.3 ps
- need picosecond-accurate time-stamping → usually high-end electronics, but also done with ASICs, FPGAs

(Single-photon) Avalanche Photodiodes

Linear mode (i.e., avalanche photodiode or *APD*):

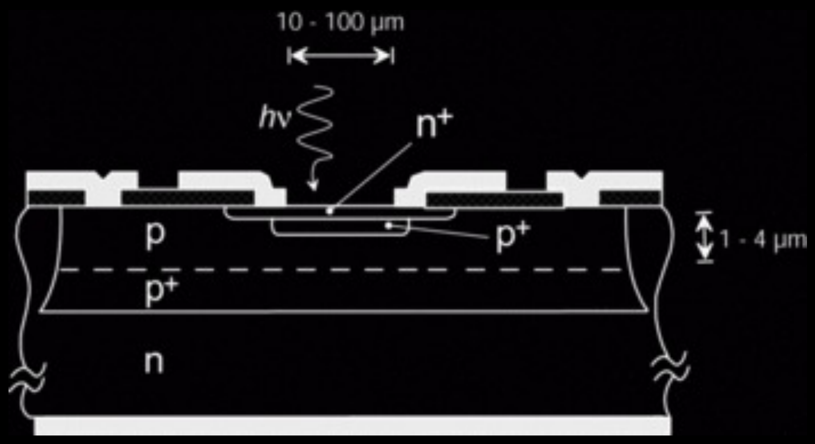
acts like a conventional photodiode with extremely high gain or amplification

time resolution >300 ps – 10 ns

Geiger mode (i.e., single-photon avalanche photodiode *SPAD*):

500x more sensitive, i.e. single-photon sensitive

time resolution ~50 ps



Semiconductor devices

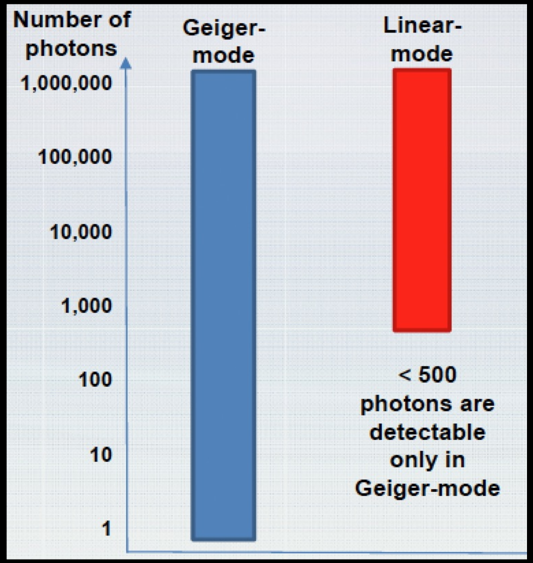
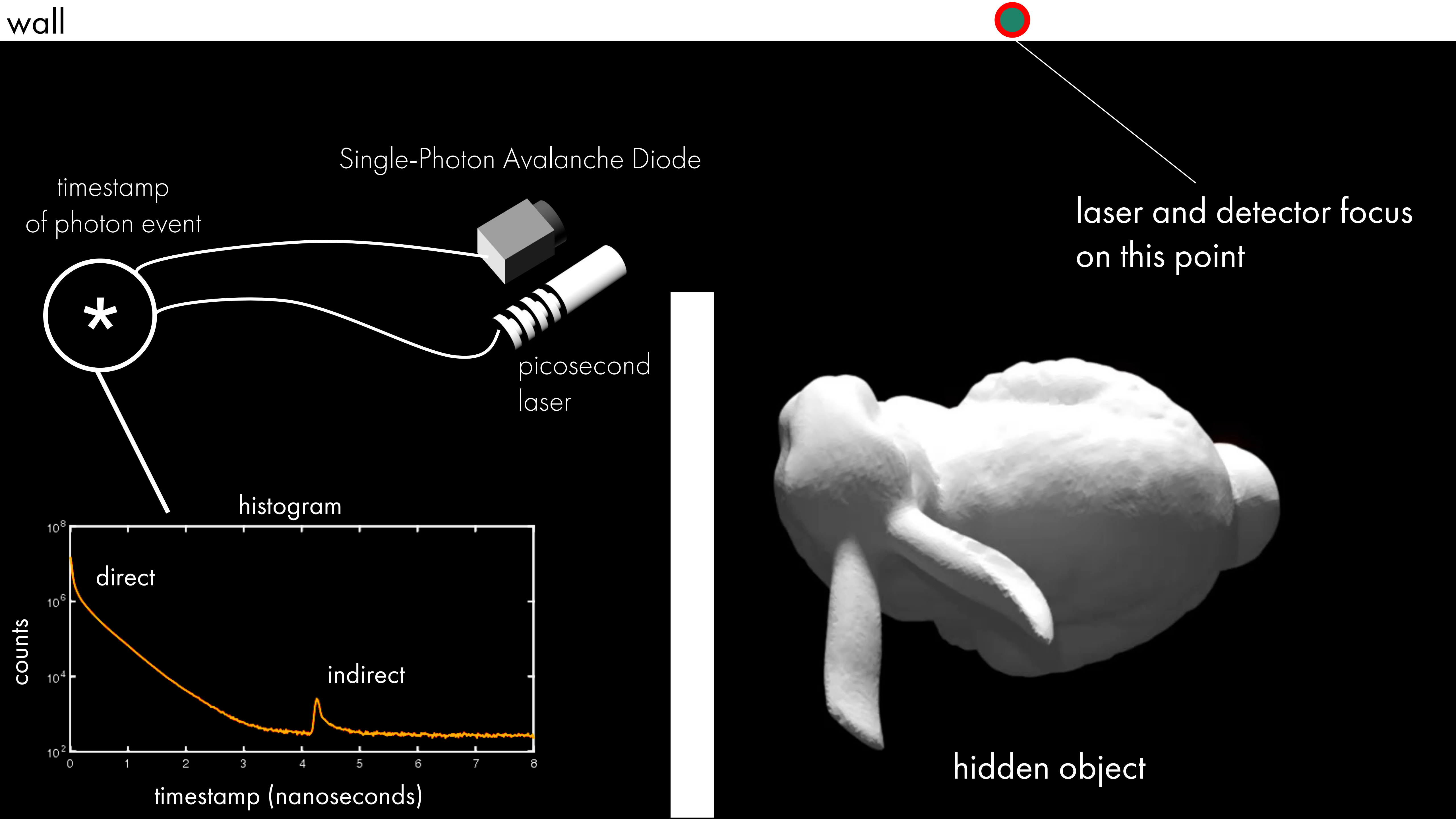
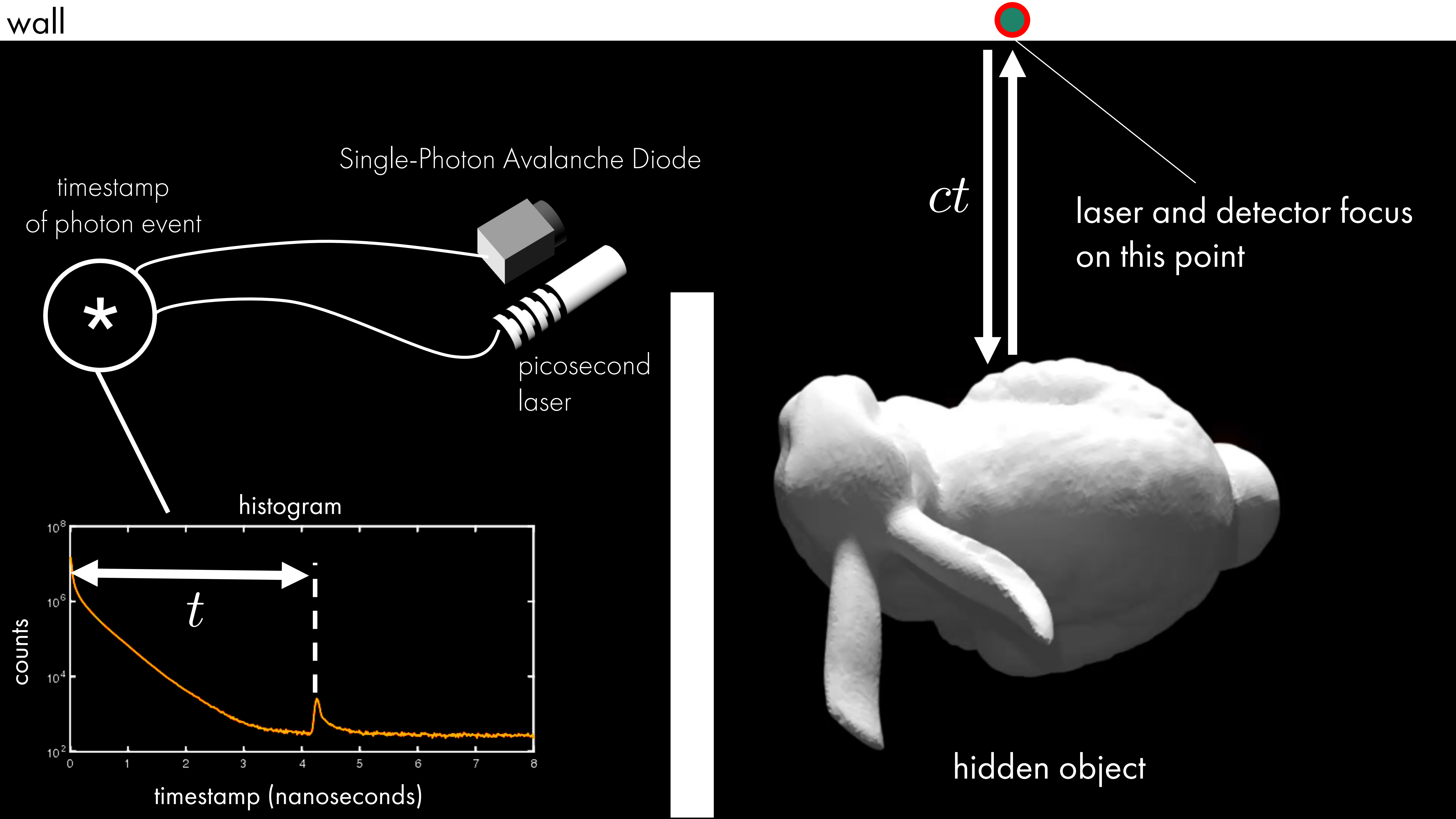


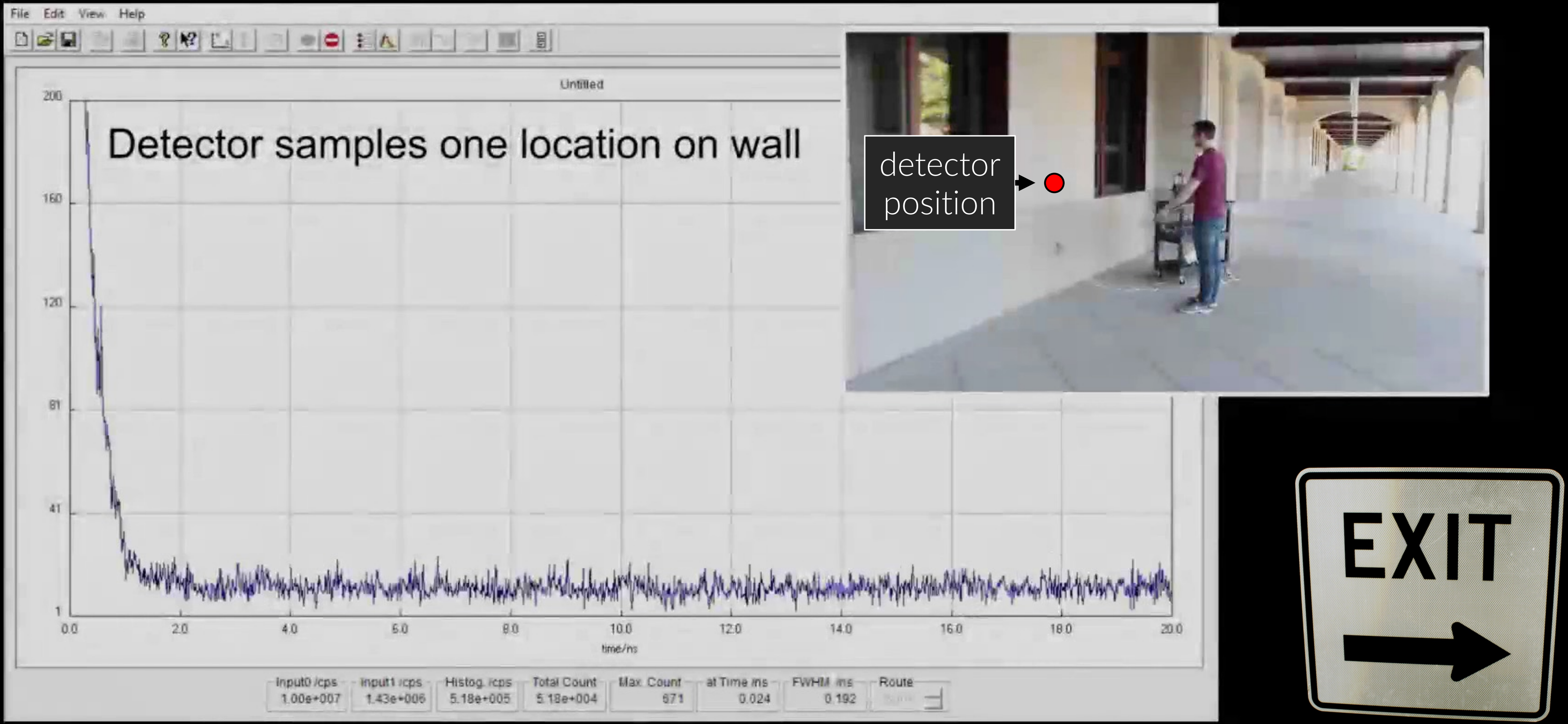
image by Princeton Lightwave


wall
laser and detector focus on this point
Single-Photon Avalanche Diode
timestamp of photon event
*
picosecond laser
histogram
direct
indirect
counts
timestamp (nanoseconds)
10^8
10^6
10^4
10^2
0 1 2 3 4 5 6 7 8
hidden object



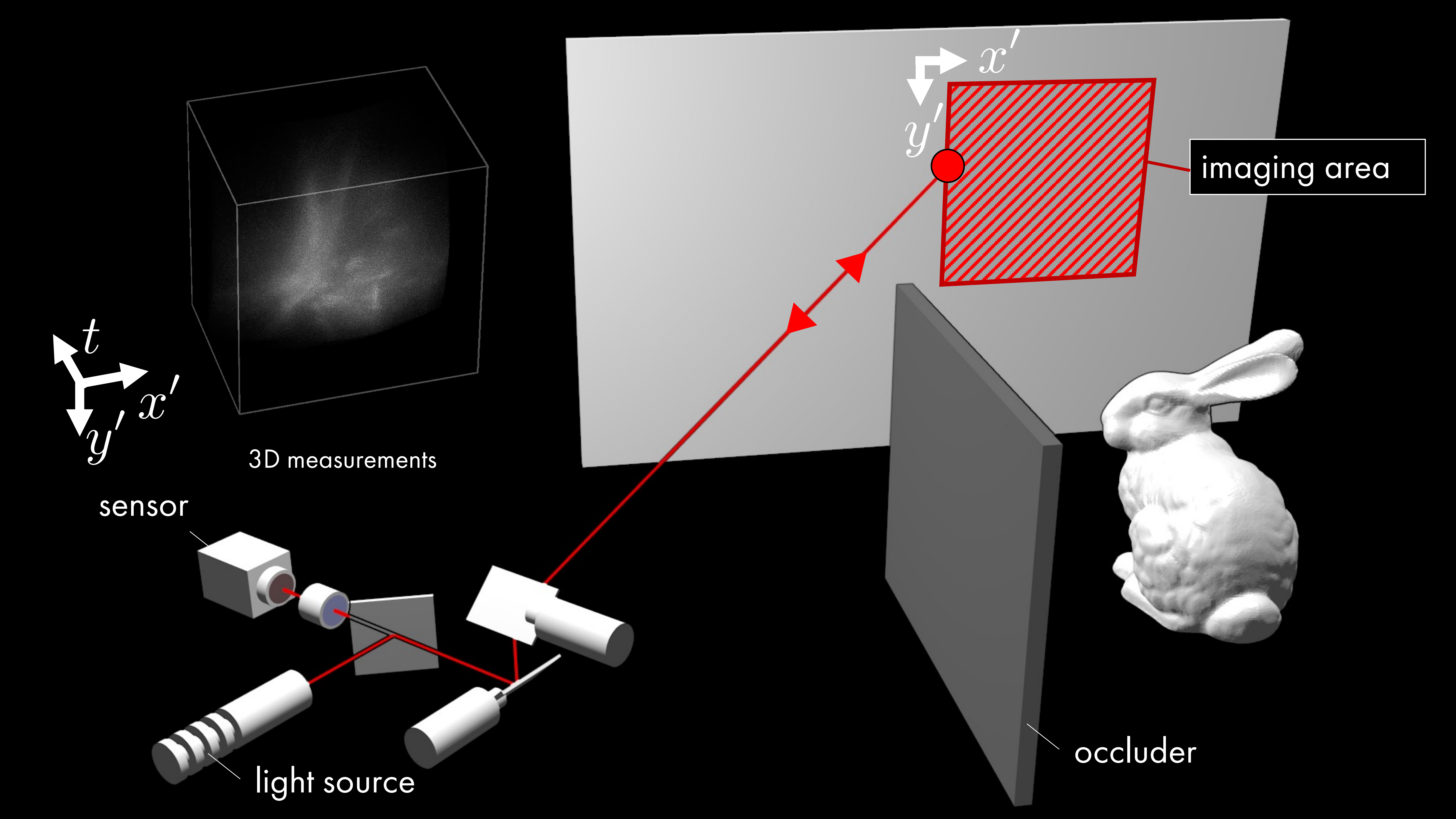
wall
Single-Photon Avalanche Diode
timestamp
of photon event
*
picosecond
laser
ct
laser and detector focus
on this point
histogram
t
counts
timestamp (nanoseconds)
hidden object

RAW histogram (10 FPS)

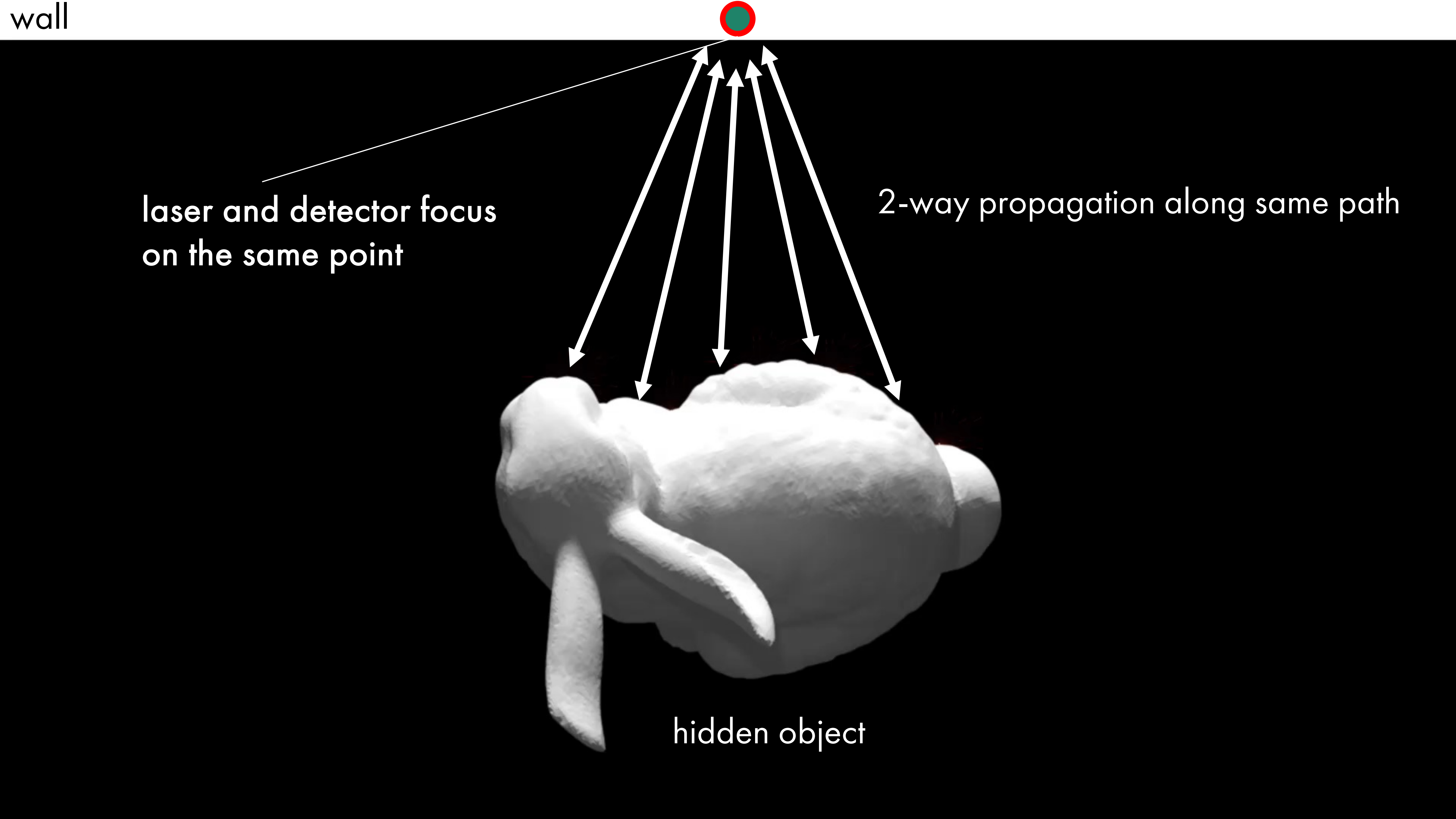


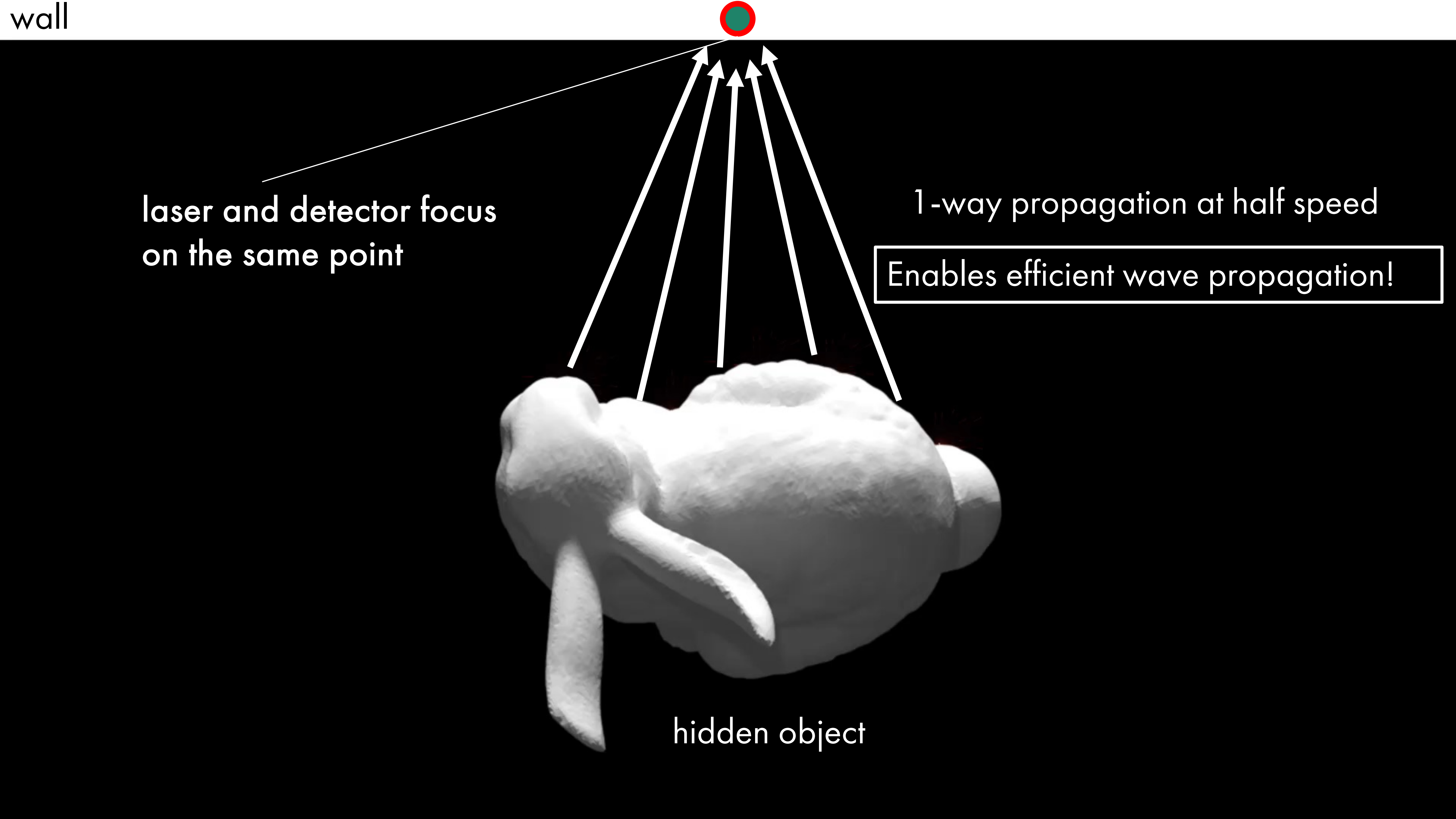
object

O'Toole et al. 2018



x'
y'
imaging area
t
x'
y'
3D measurements
sensor
occluder
light source


wall
laser and detector focus
on the same point
2-way propagation along same path
hidden object


wall
laser and detector focus on the same point
1-way propagation at half speed
Enables efficient wave propagation!
hidden object

$$\nabla^2 \Psi - \frac{1}{v^2} \frac{\partial^2 \Psi}{\partial t^2} = 0$$

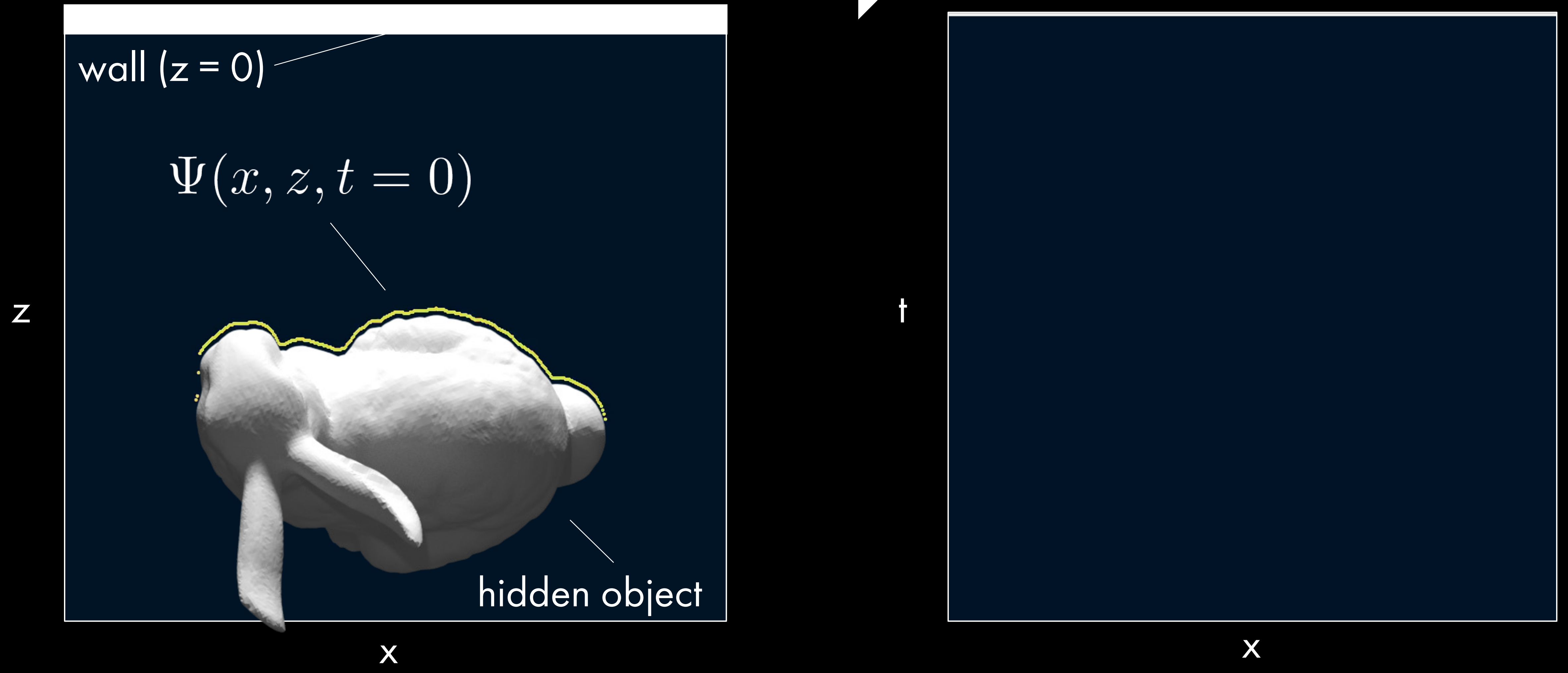
image formation model

$$\Psi(x, z, t)$$

wavefield

$$\Psi(x, z = 0, t)$$

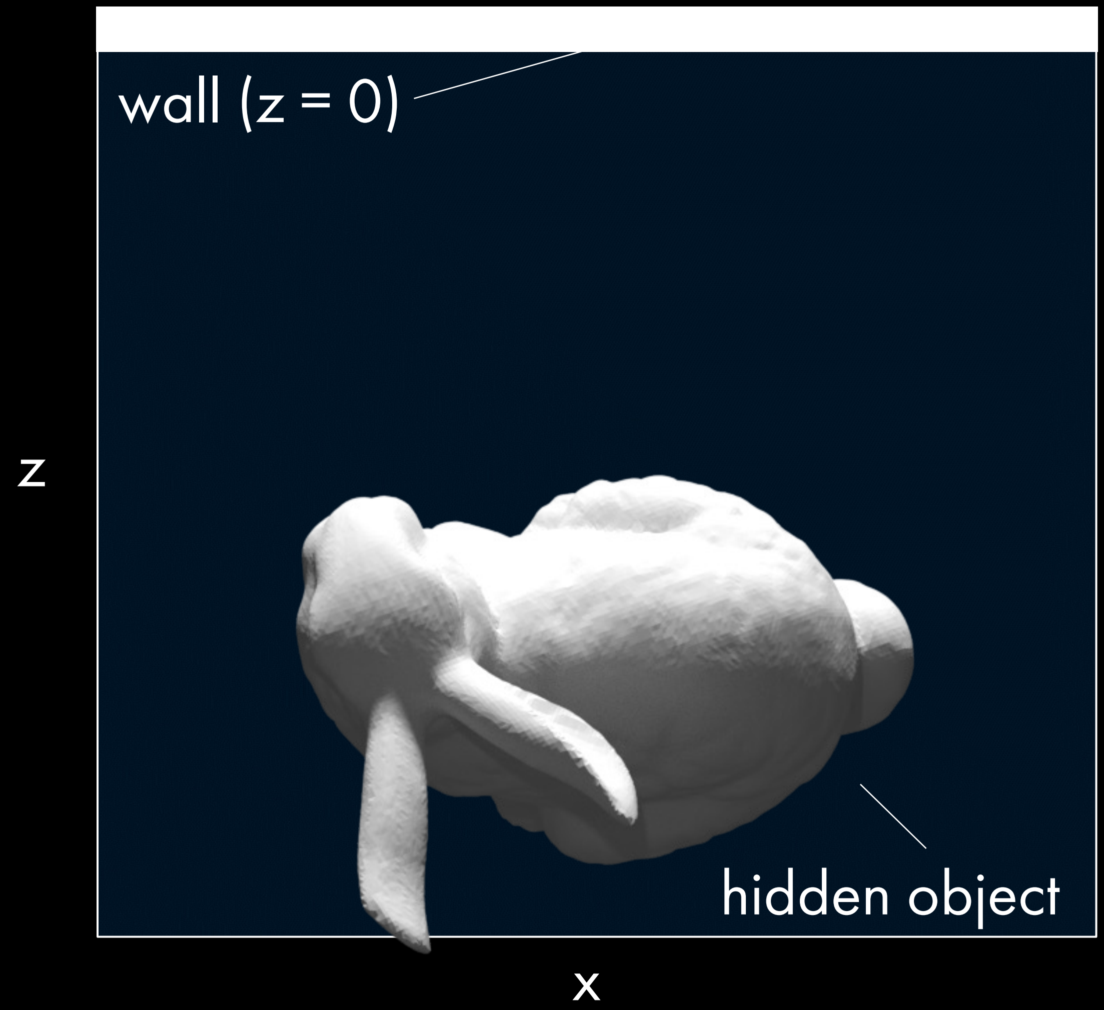
confocal measurements



general solution (time reversal)

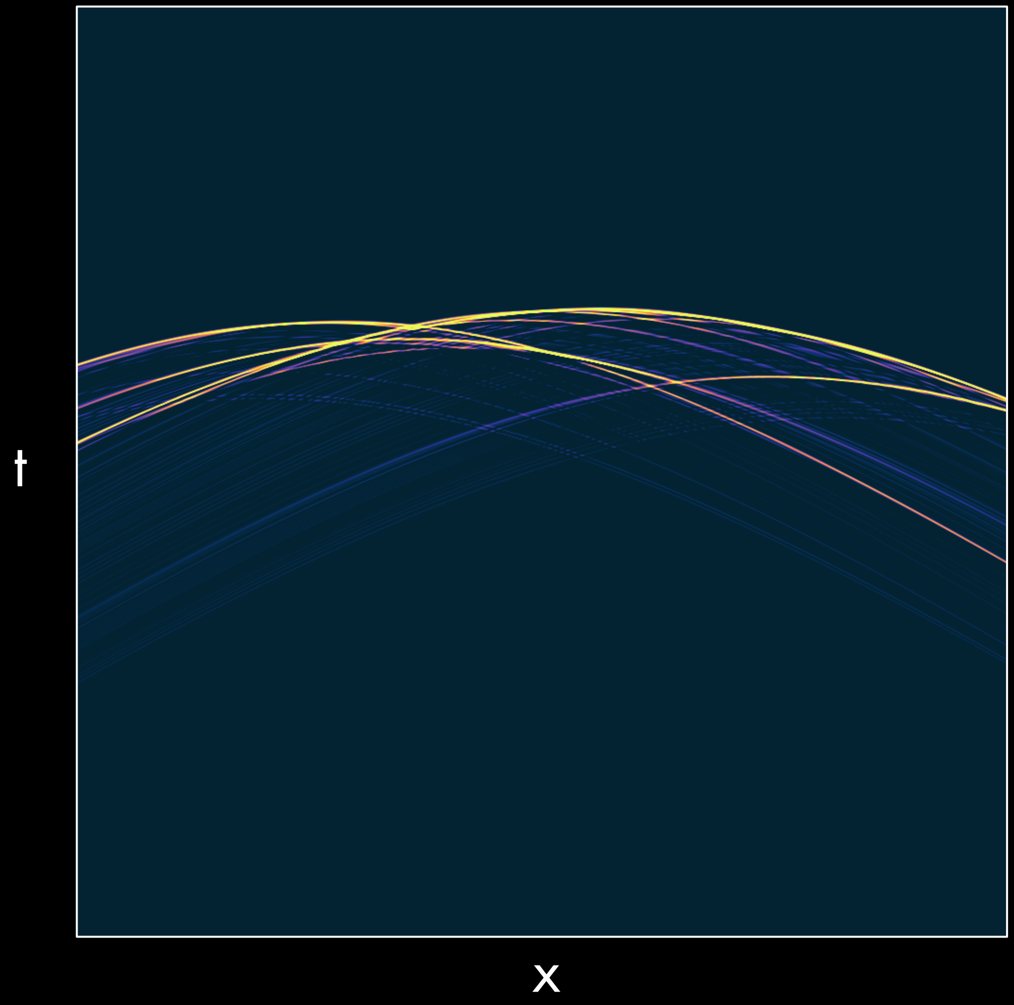
$$\Psi(x, z, t = \textbf{03.000 ns})$$

wavefield



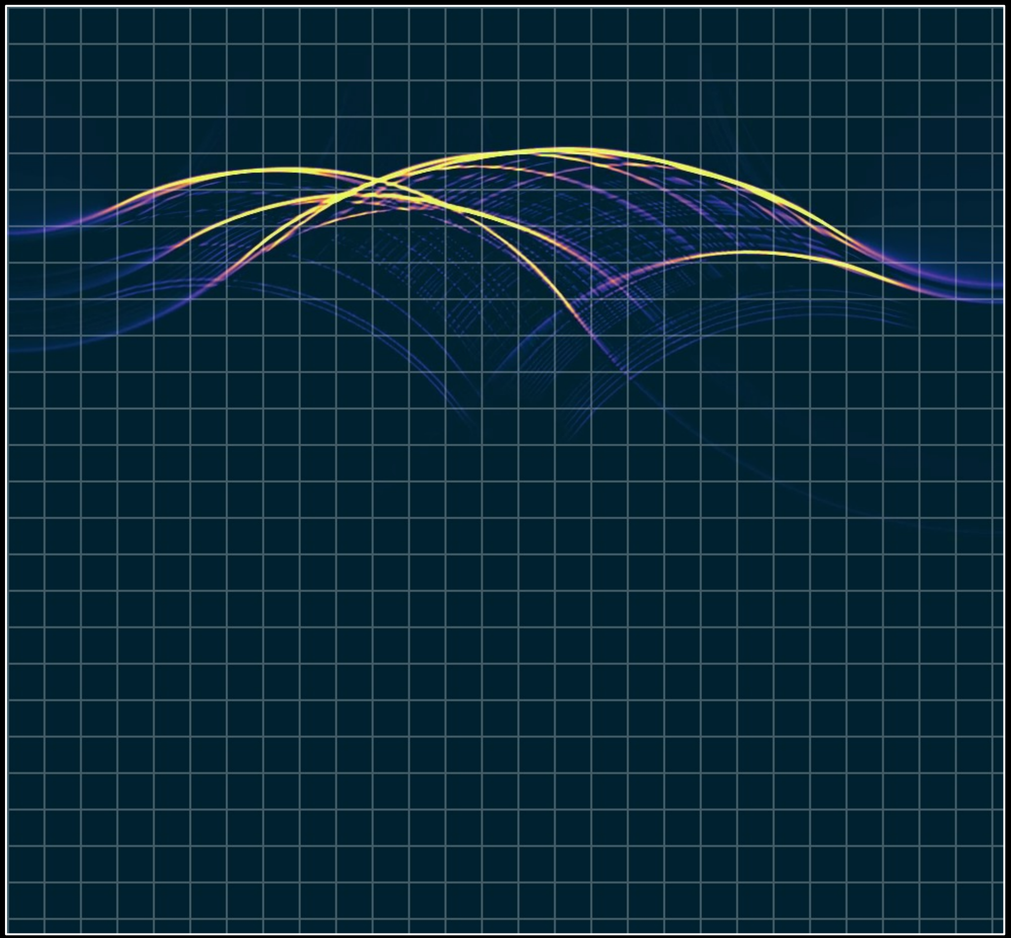
$$\Psi(x, z = 0, t)$$

confocal measurements



general solution (time reversal)

finite-difference time-domain method



1. approximate wave equation with finite differences

$$\frac{\partial^2 \Psi}{\partial t^2} \approx \frac{\Psi_i^{n+1} - 2\Psi_i^n + \Psi_i^{n-1}}{(\Delta t)^2}$$

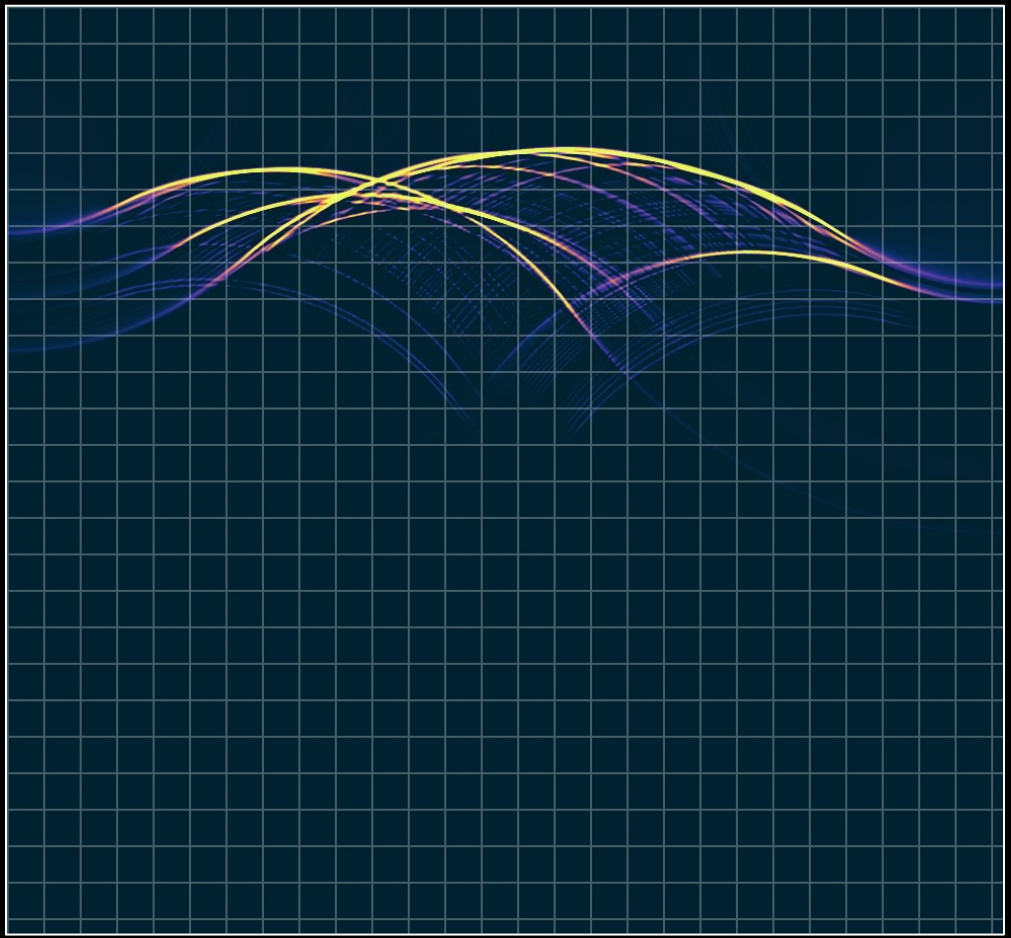
2. solve for previous timestep

$$\Psi_i^{n-1} = f\left(\Psi^n, \Psi^{n+1}\right)$$

3. repeatedly update Ψ at all grid cells

general solution (time reversal)

finite-difference time-domain method



1. approximate wave equation with finite differences

$$\frac{\partial^2 \Psi}{\partial t^2} \approx \frac{\Psi_i^{n+1} - \ldots\, {}^{n-1}}{\ldots}$$

Slow to get t=0 at high-resolution!

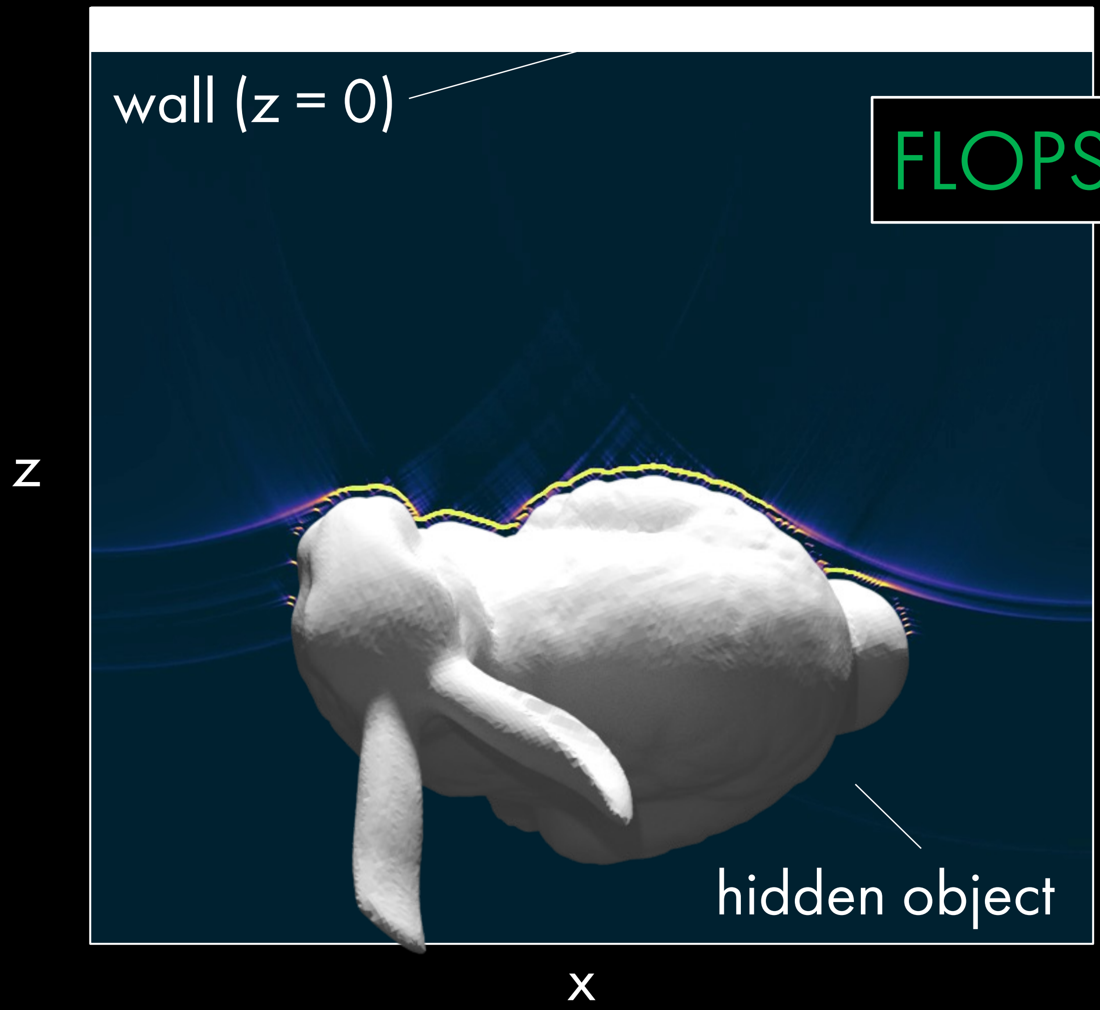
2. solve f… timestep

$$\Psi_i^{n-1} = f\left(\Psi^n, \Psi^{n+1}\right)$$

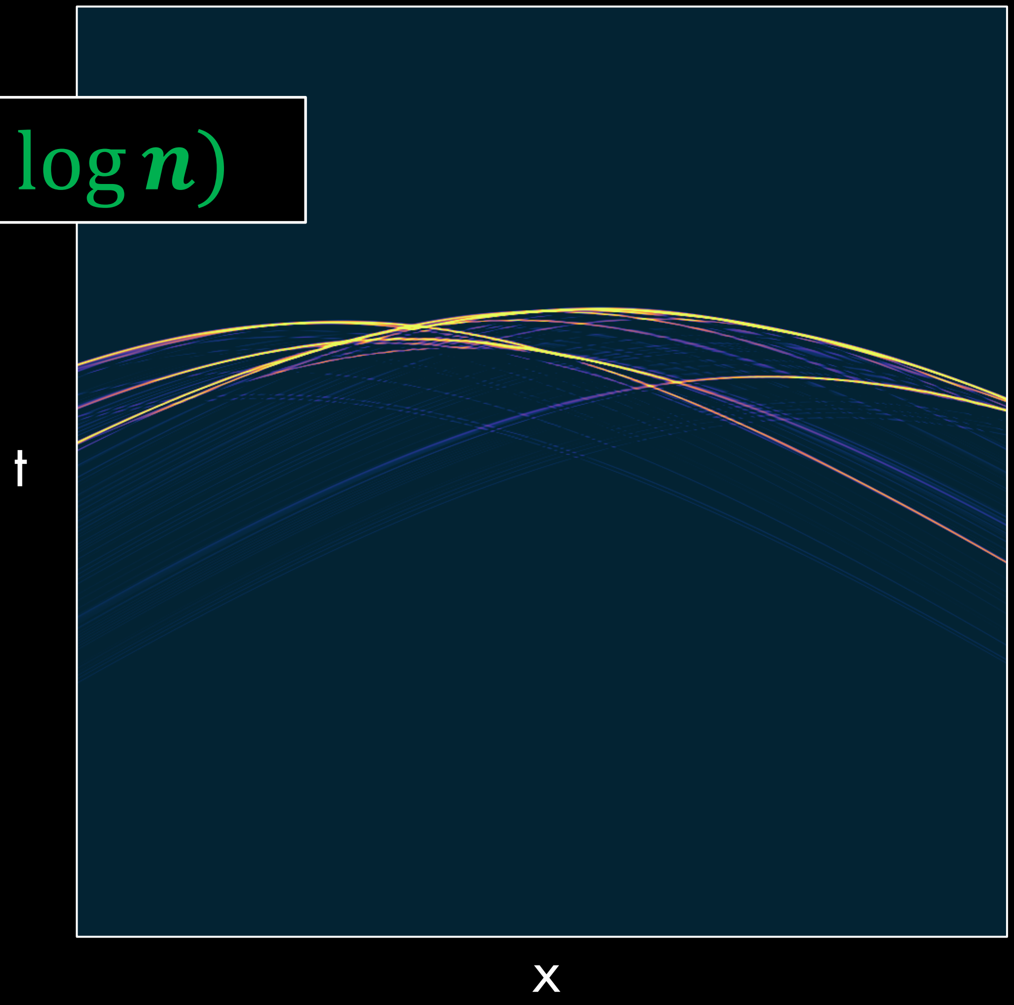
3. repeatedly update Ψ at all grid cells

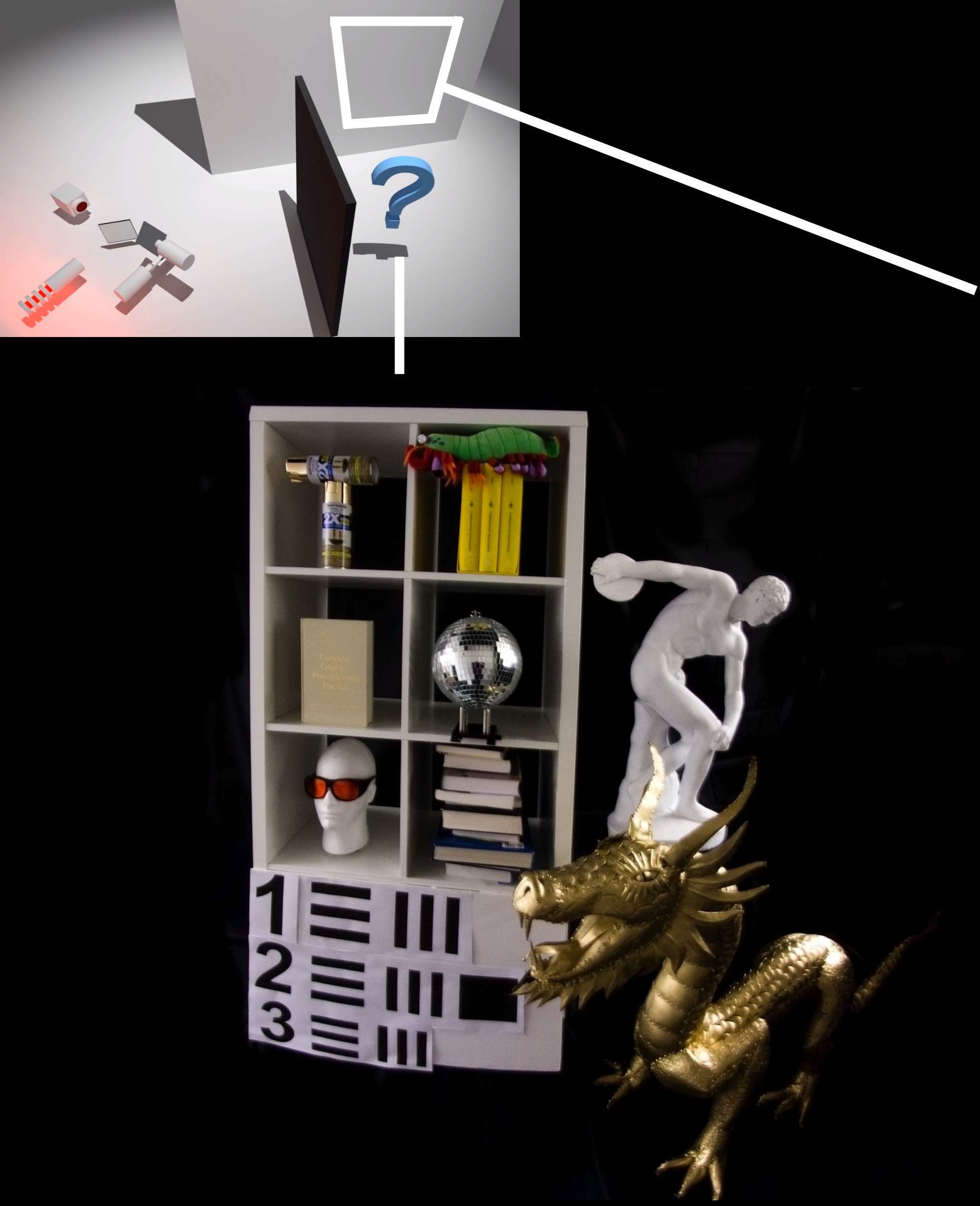
frequency–wavenumber (*f*–*k*) Migration

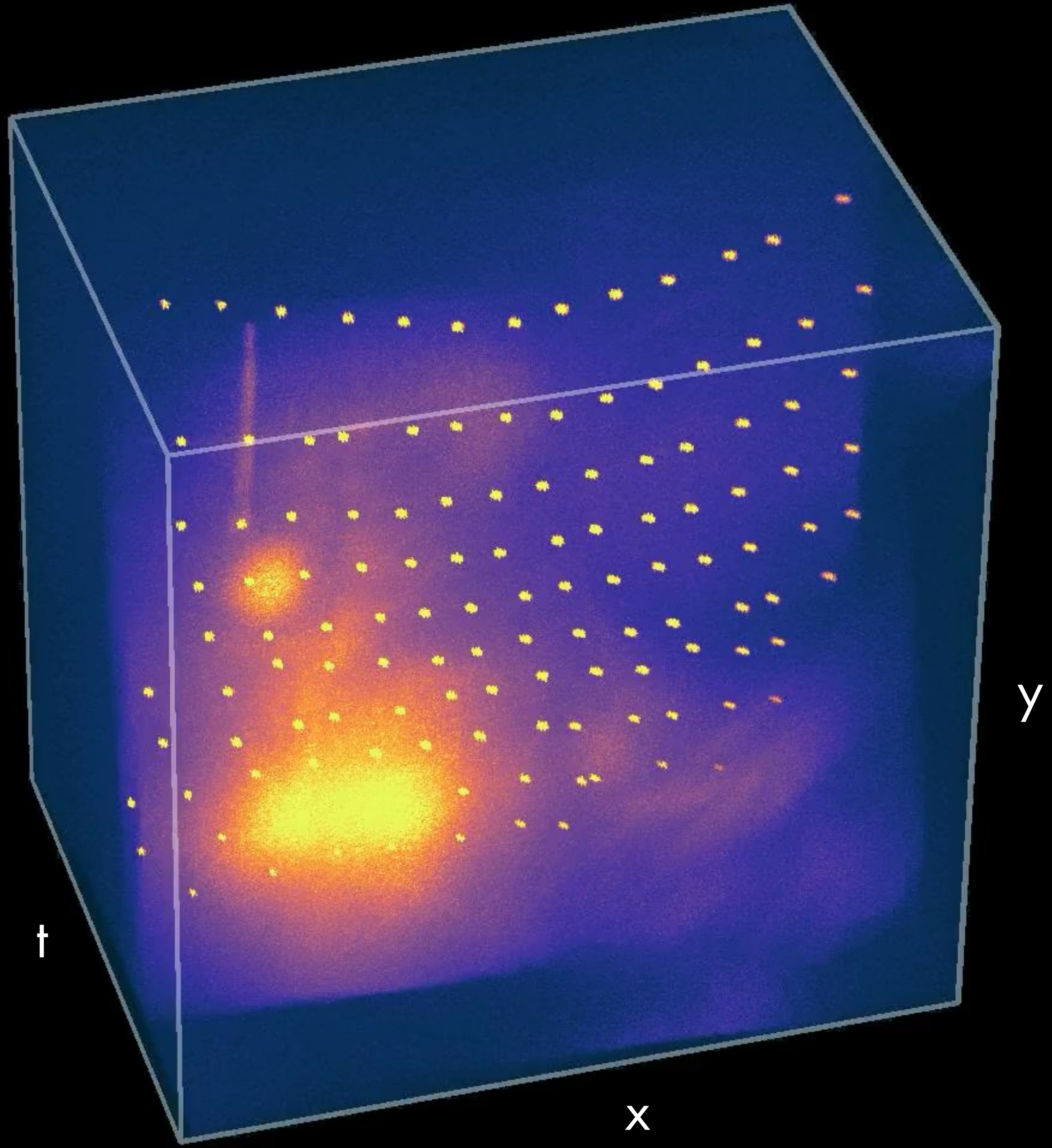
$\Psi(x, z, t = 0)$ wavefield ← $\Psi(x, z = 0, t)$ confocal measurements



FLOPS: $O(n^3 \log n)$

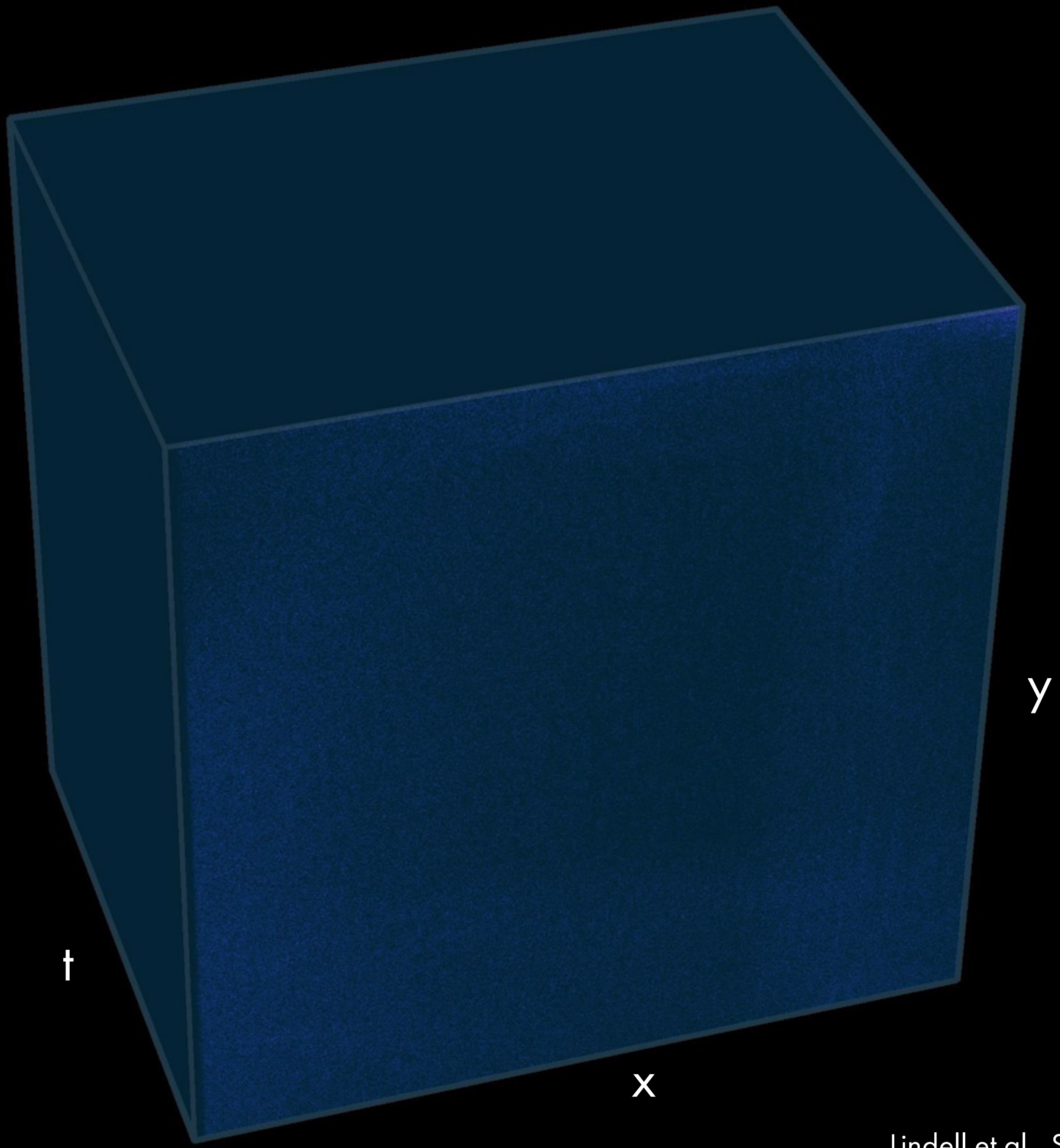






Captured Measurements

Lindell et al., SIGGRAPH 2019

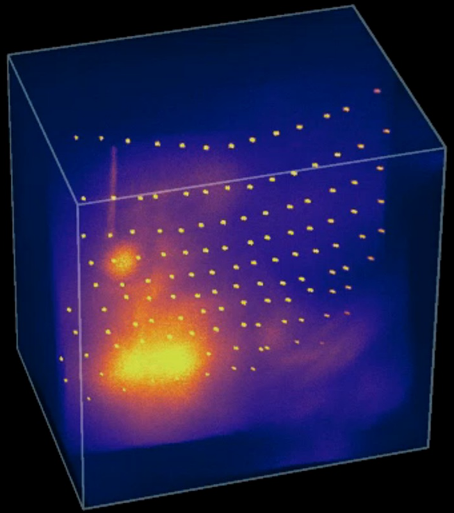


Lindell et al., SIGGRAPH 2019

f–k Migration

$\Psi(x, y, t)$

Measurements (z=0)



$\bar{\Phi}(k_x, k_y, f)$

Spectrum

$\Phi(k_x, k_y, k_z)$

Interpolated Spectrum

$\Psi(x, y, z)$

Hidden Volume (t=0)

$\mathcal{F}$

Resample

$\mathcal{F}^{-1}$

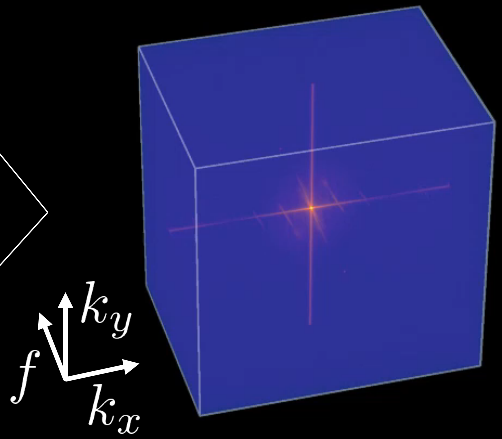
f–k Migration

Express wavefield as function of measurement spectrum (plane wave decomposition)

$$\Psi(x,y,z,t) = \iiint \bar{\Phi}(k_x,k_y,f)e^{2\pi i(k_x x + k_y y + k_z z - ft)}dk_x dk_y df$$

wavefield

Fourier transform of measurements



Set t=0 to get migrated solution

$$\Psi(x,y,z,t=0) = \iiint \bar{\Phi}(k_x,k_y,f)e^{2\pi i(k_x x + k_y y + \underbrace{k_z z})}dk_x dk_y \underbrace{df}$$

Almost an inverse Fourier Transform!

f–k Migration

Set t=0 to get migrated solution

$$\Psi(x,y,z,t=0) = \iiint \bar{\Phi}(k_x,k_y,f)e^{2\pi i(k_x x + k_y y + \underbrace{k_z}z)}dk_x dk_y \underbrace{df}$$

Almost an inverse Fourier Transform!

Use dispersion relation[1] to perform substitution of variables

$$f = v\sqrt{k_x^2 + k_y^2 + k_z^2}$$

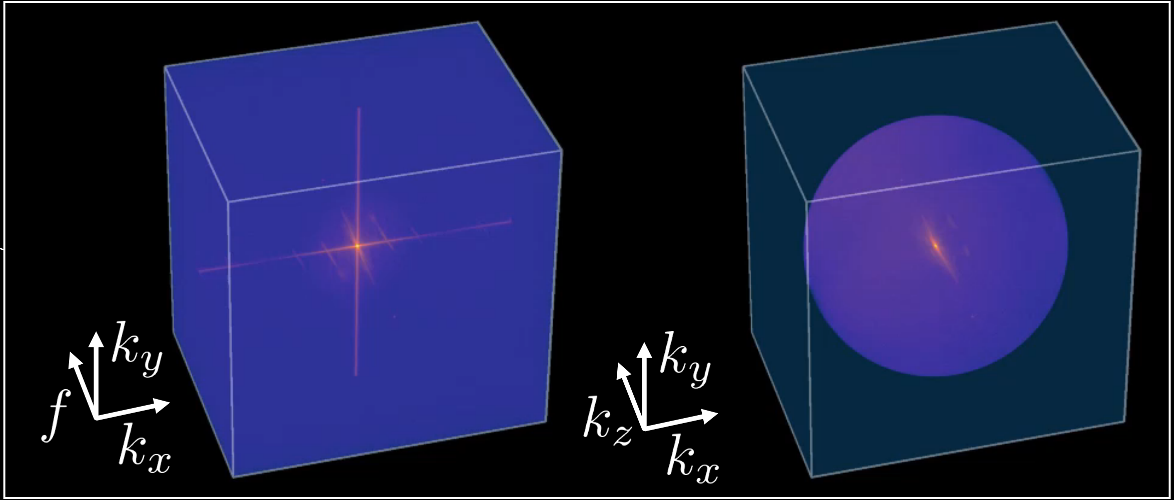
$$\boxed{f \Rightarrow k_z}$$

[1]Georgi, Howard. *The physics of waves*. Englewood Cliffs, NJ: Prentice Hall, 1993.

Use dispersion relation[1] to perform substitution of variables

$$f = v\sqrt{k_x^2 + k_y^2 + k_z^2}$$

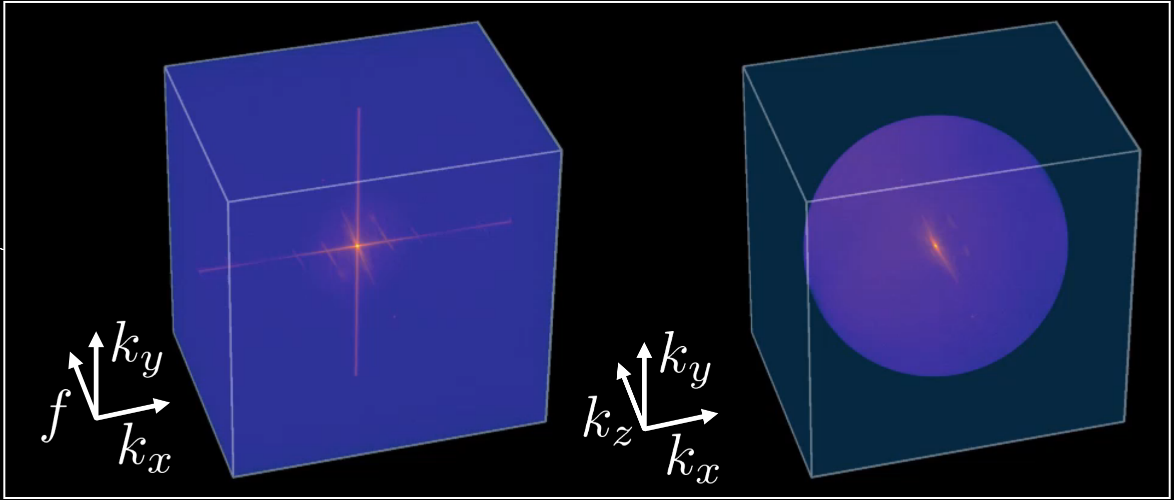
$$\boxed{f \Rightarrow k_z}$$

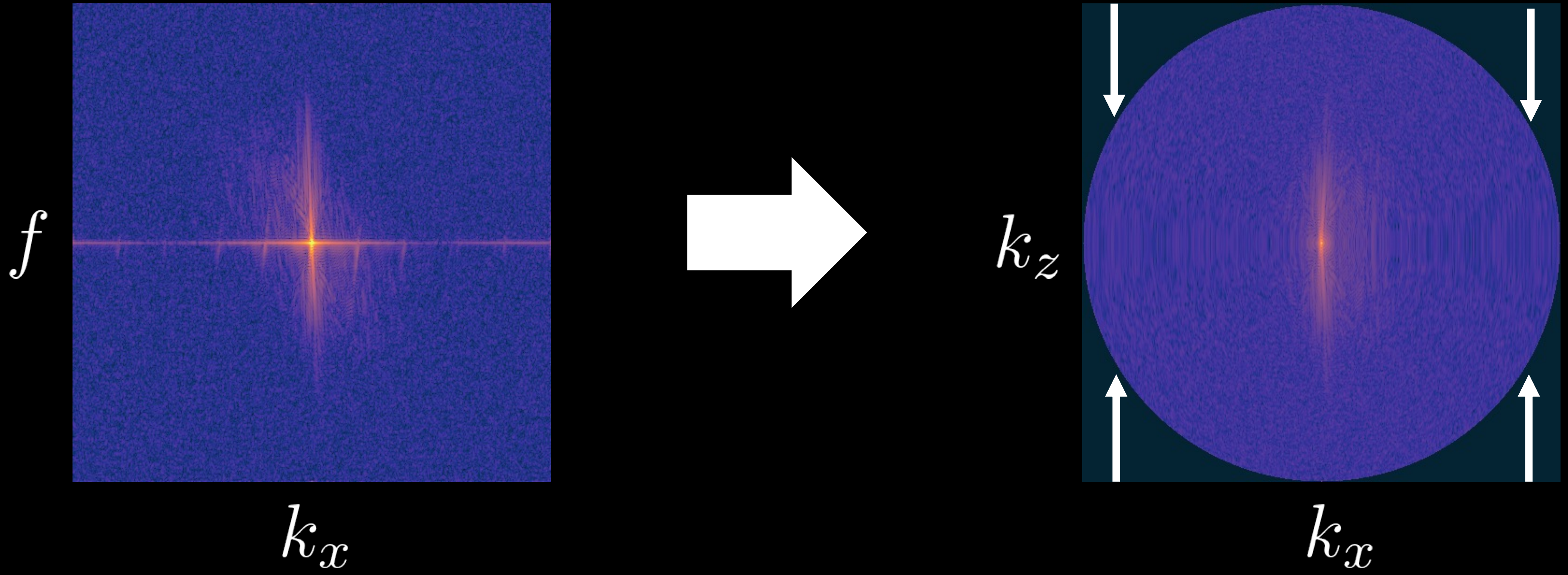


Use dispersion relation[1] to perform substitution of variables

$$f = v\sqrt{k_x^2 + k_y^2 + k_z^2}$$

$$\boxed{f \Rightarrow k_z}$$

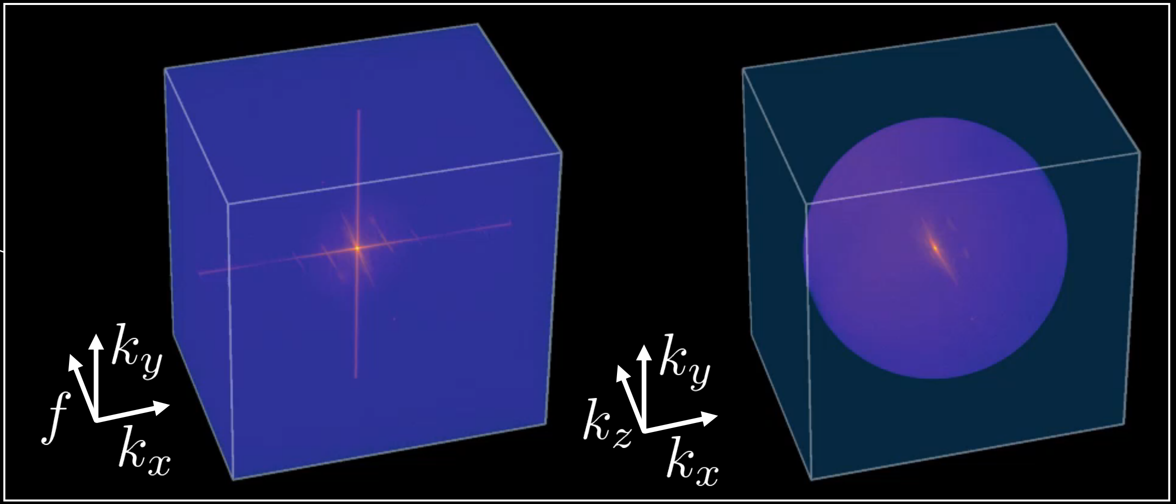




Use dispersion relation[1] to perform substitution of variables

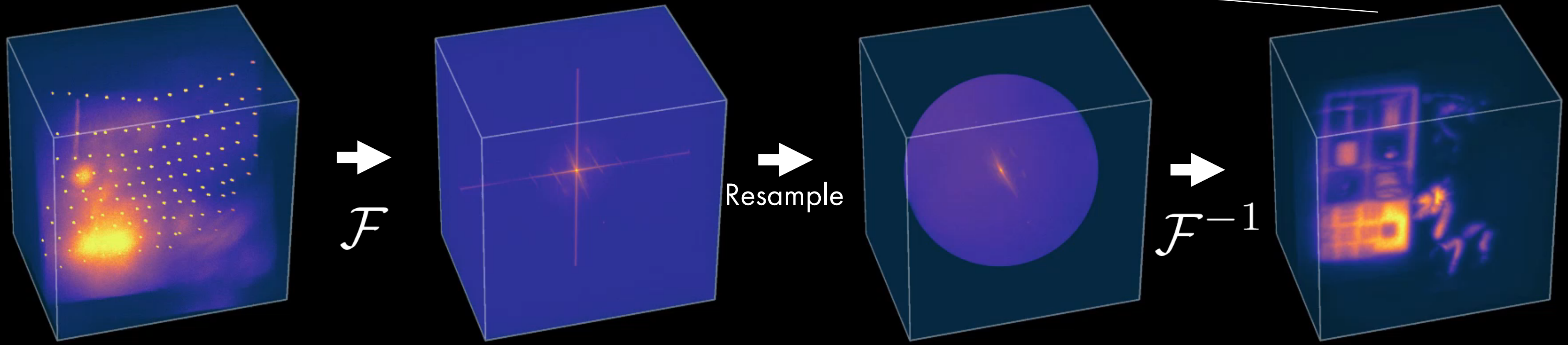
$$f = v\sqrt{k_x^2 + k_y^2 + k_z^2}$$

$$\boxed{f \Rightarrow k_z}$$

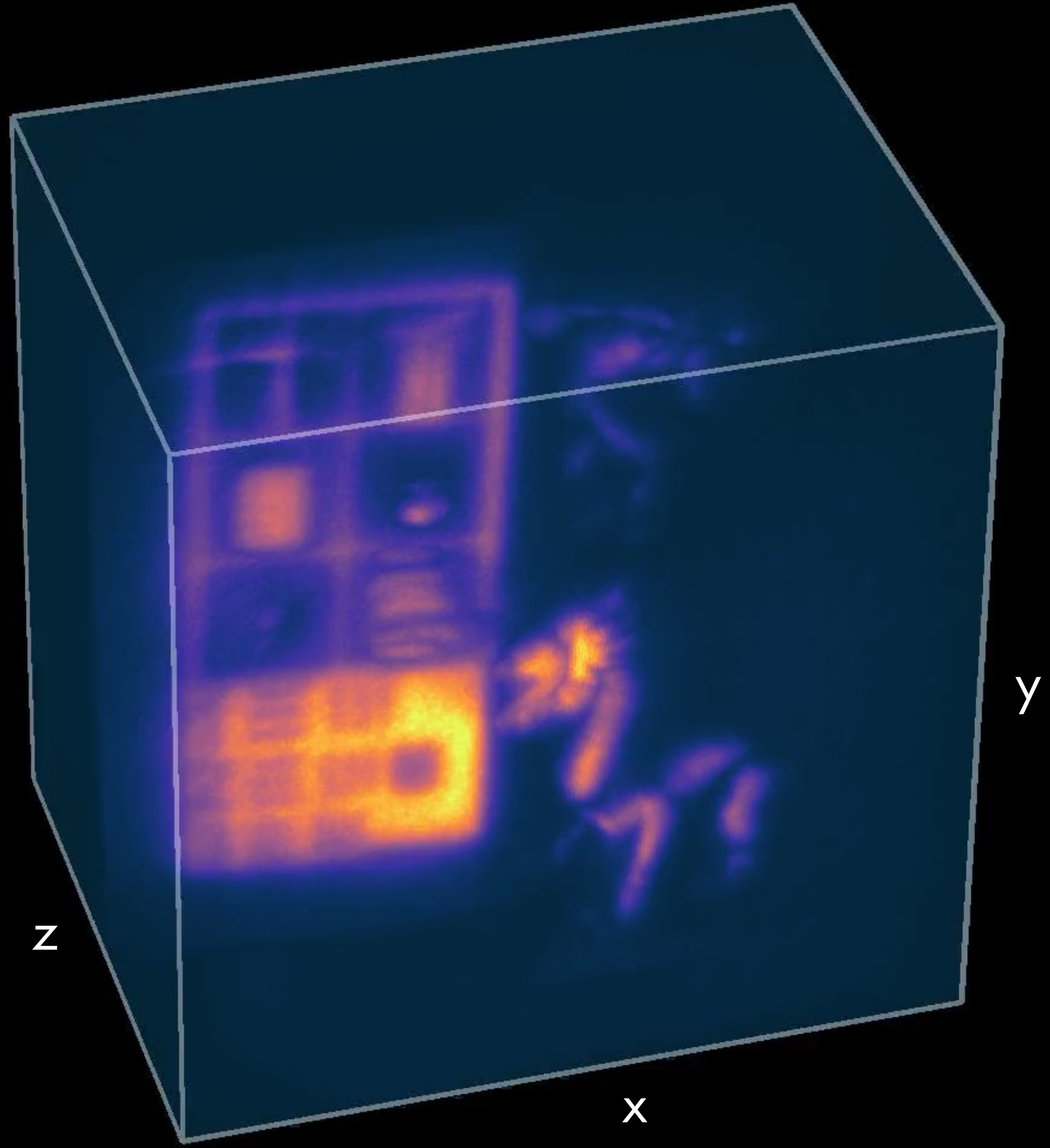


The migrated solution is an inverse Fourier Transform!

$$\Psi(x, y, z, t = 0) = \iiint \Phi(k_x, k_y, k_z) e^{2\pi i (k_x x + k_y y + k_z z)} dk_x dk_y dk_z$$



f–k Migration

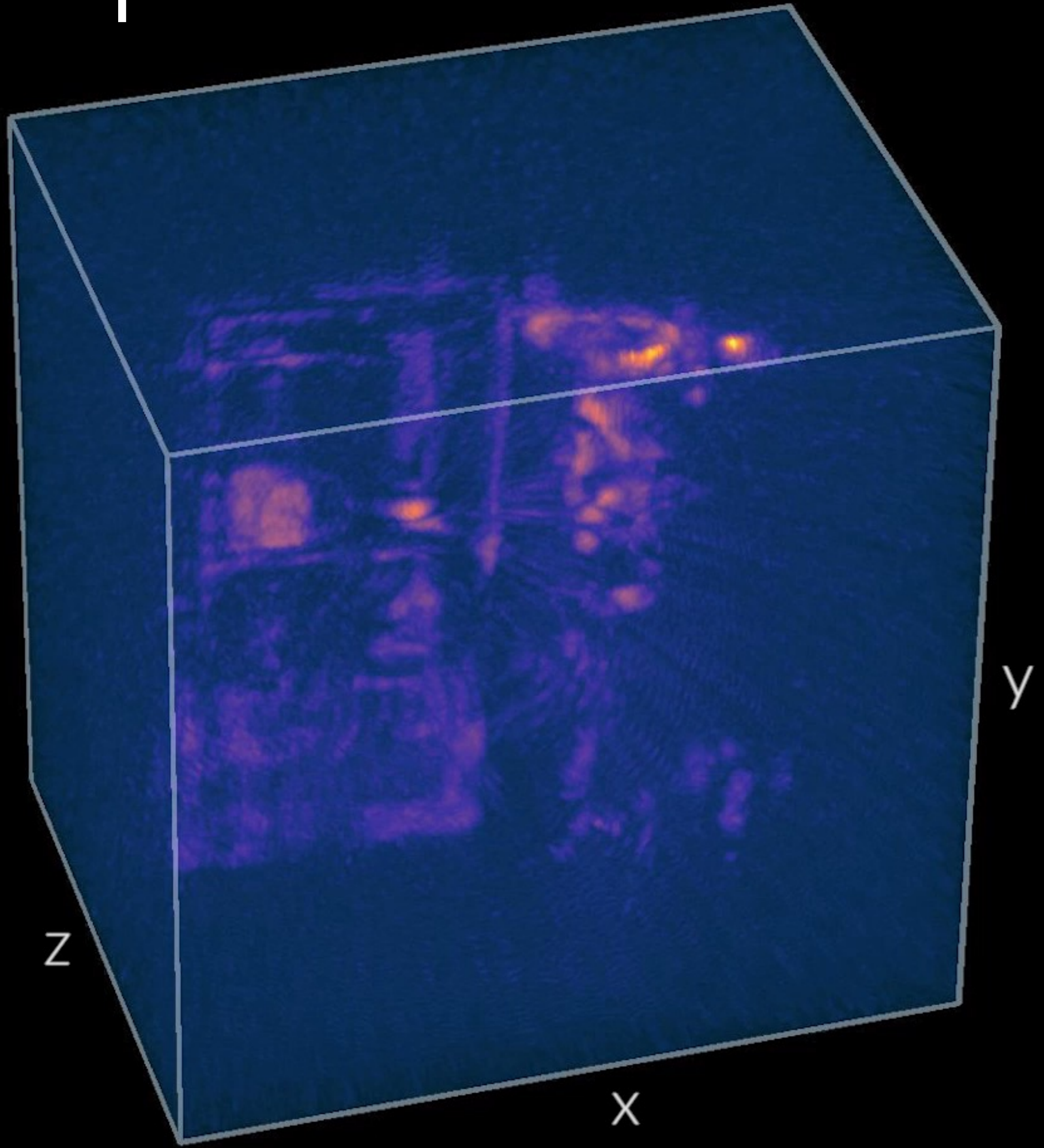


Dimensions: 2 x 2 m
Exposure: 180 min
Reconstruction time: ~90 sec (CPU)

Lindell et al., SIGGRAPH 2019

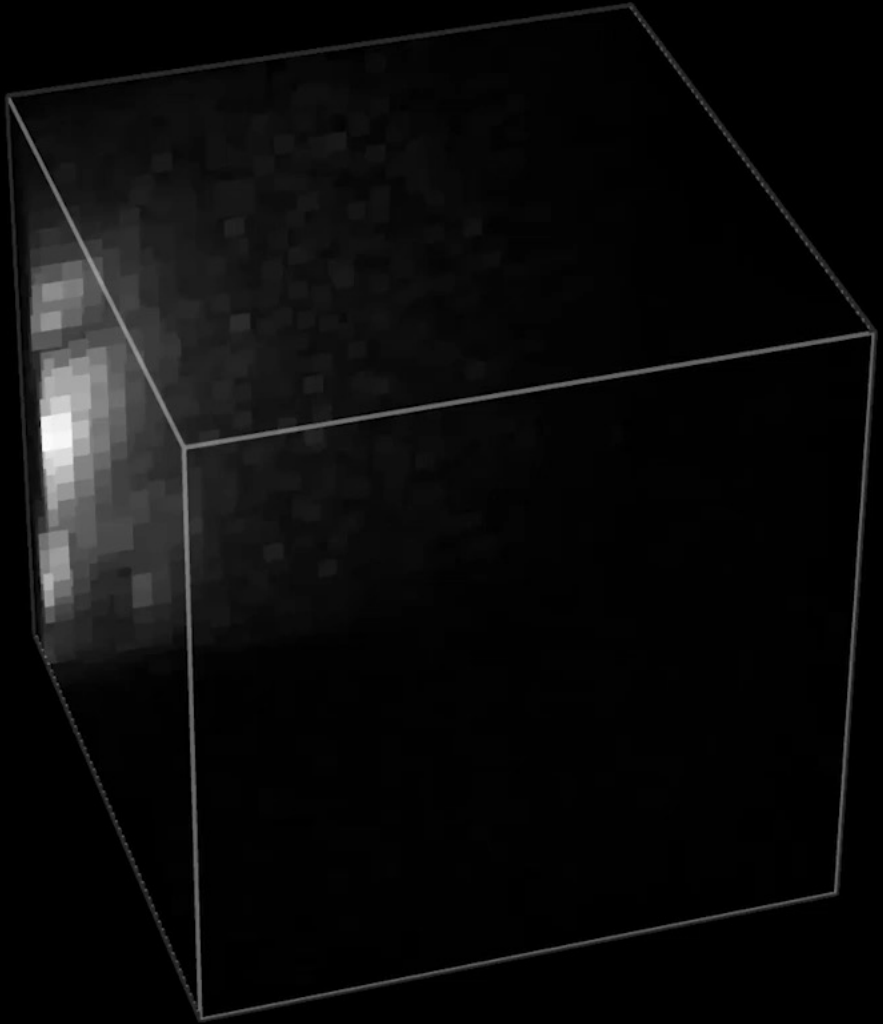
Reconstruction Comparison

dimensions: 2 m x 2 m x 1.5 m



Filtered Backprojection

real-time scanning



Framerate: 4 Hz
Resolution: 32 x 32
Dimensions: 2 m x 2 m x 2 m
Reconstruction time: ~1 s per frame

Lindell et al., SIGGRAPH 2019

Outlook

Directional Light-Cone Transform

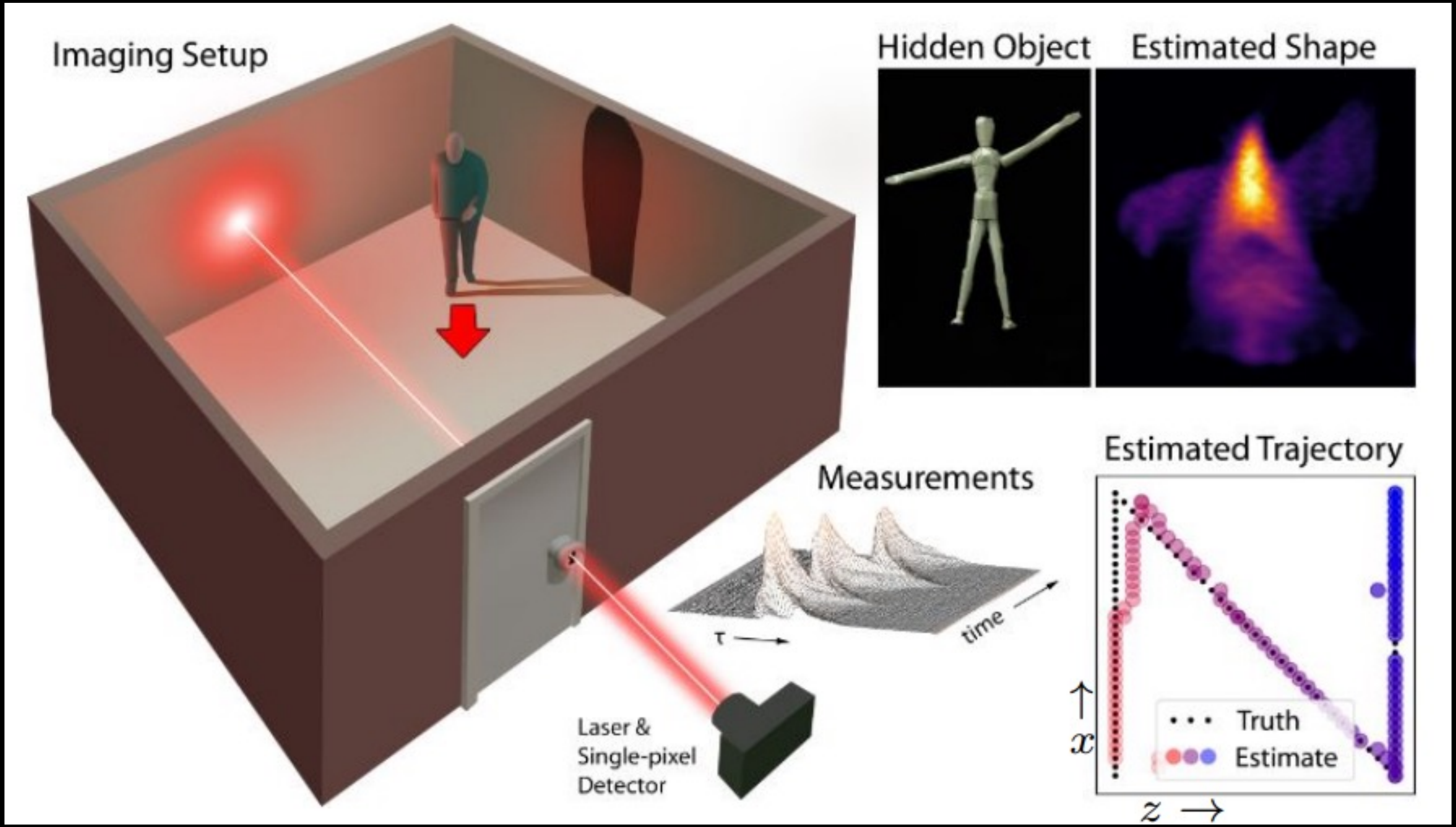
hidden scene

Recovered surface

[Young et al., CVPR 2020]

Outlook

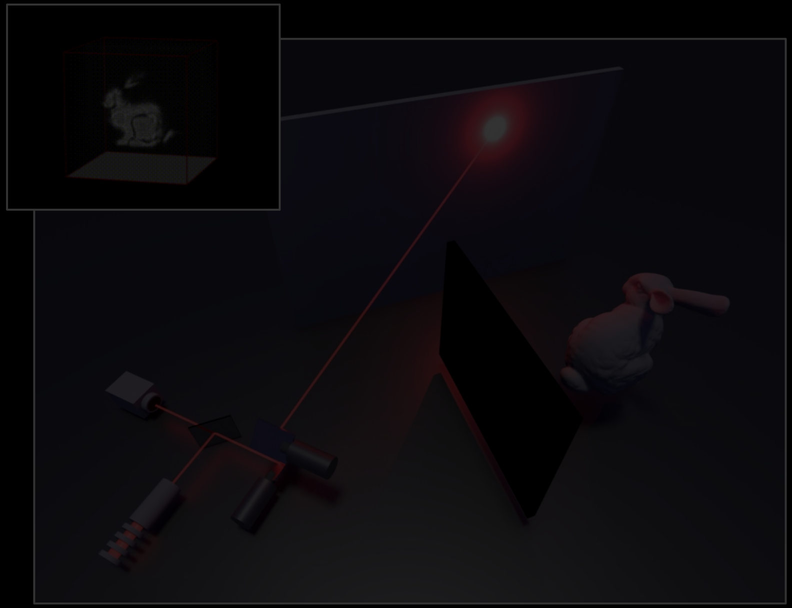
Keyhole NLOS Imaging



[Metzler et al., IEEE TCI 2021]

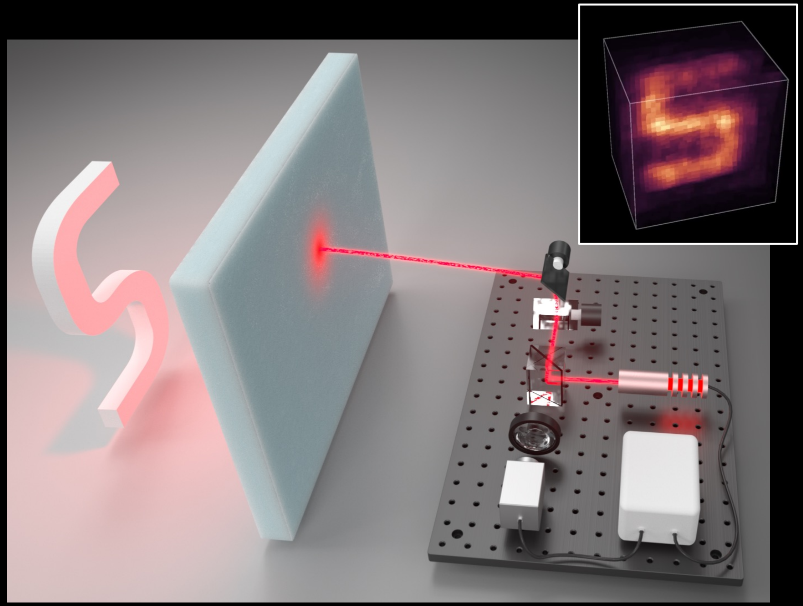
Overview

Non-Line-of-Sight Imaging



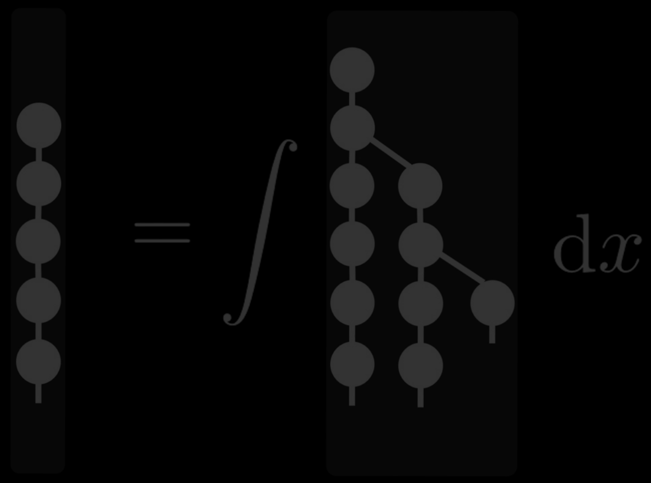
Nature '18
SIGGRAPH '19
CVPR '19
ACM Trans. Graph. '20
CVPR '20
IEEE TCI '21

Imaging through Scattering Media



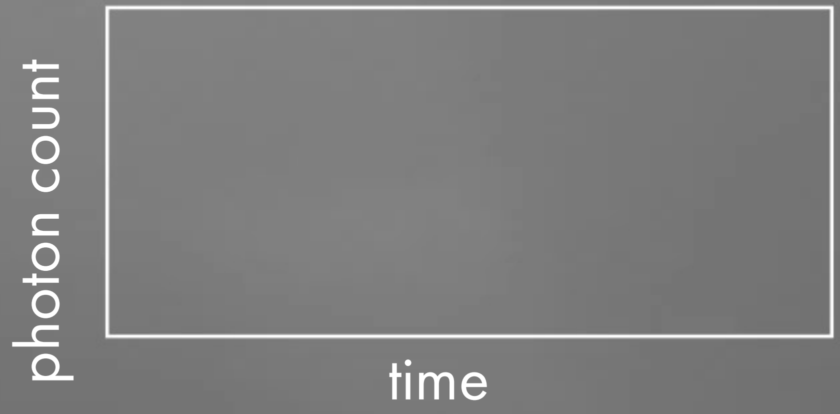
Nature Communications '20

Physics-based AI & Neural Rendering



NeurIPS '20
CVPR '21
In submission '21

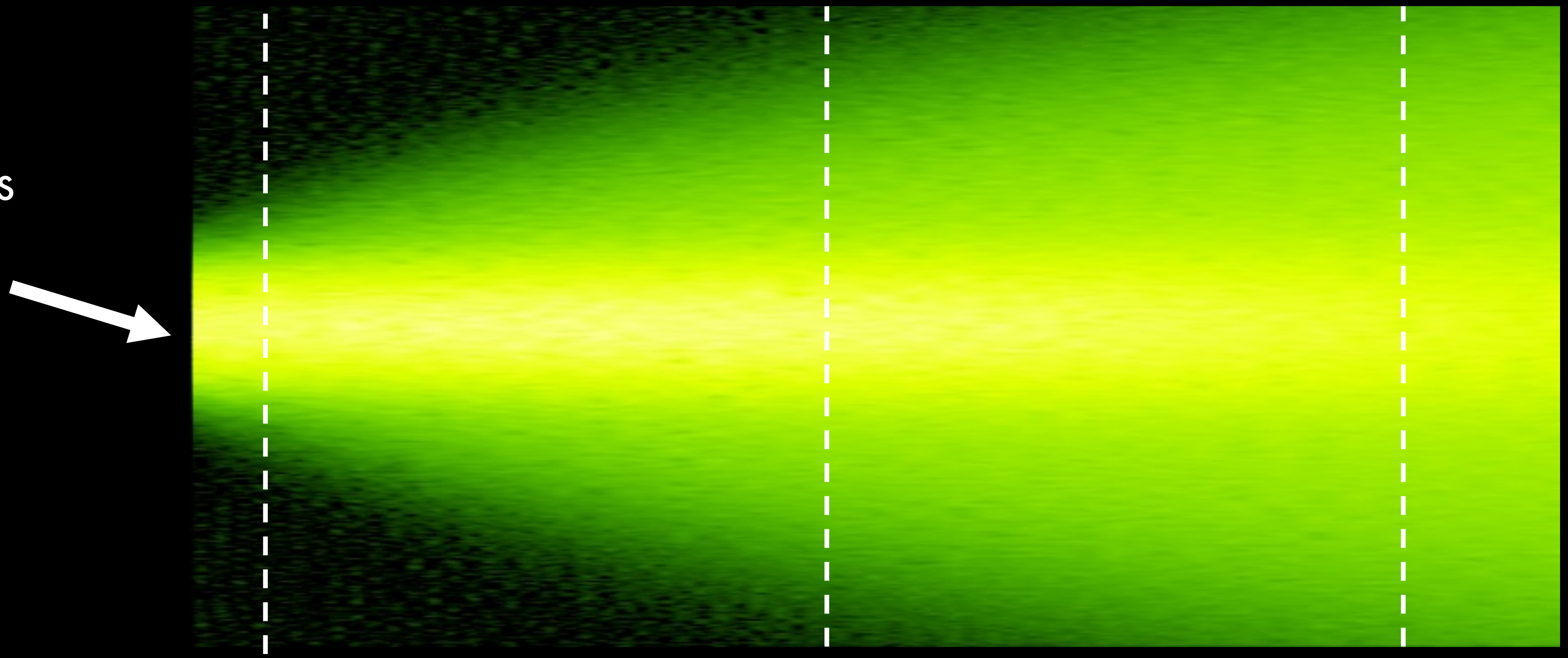
Time-resolved active imaging



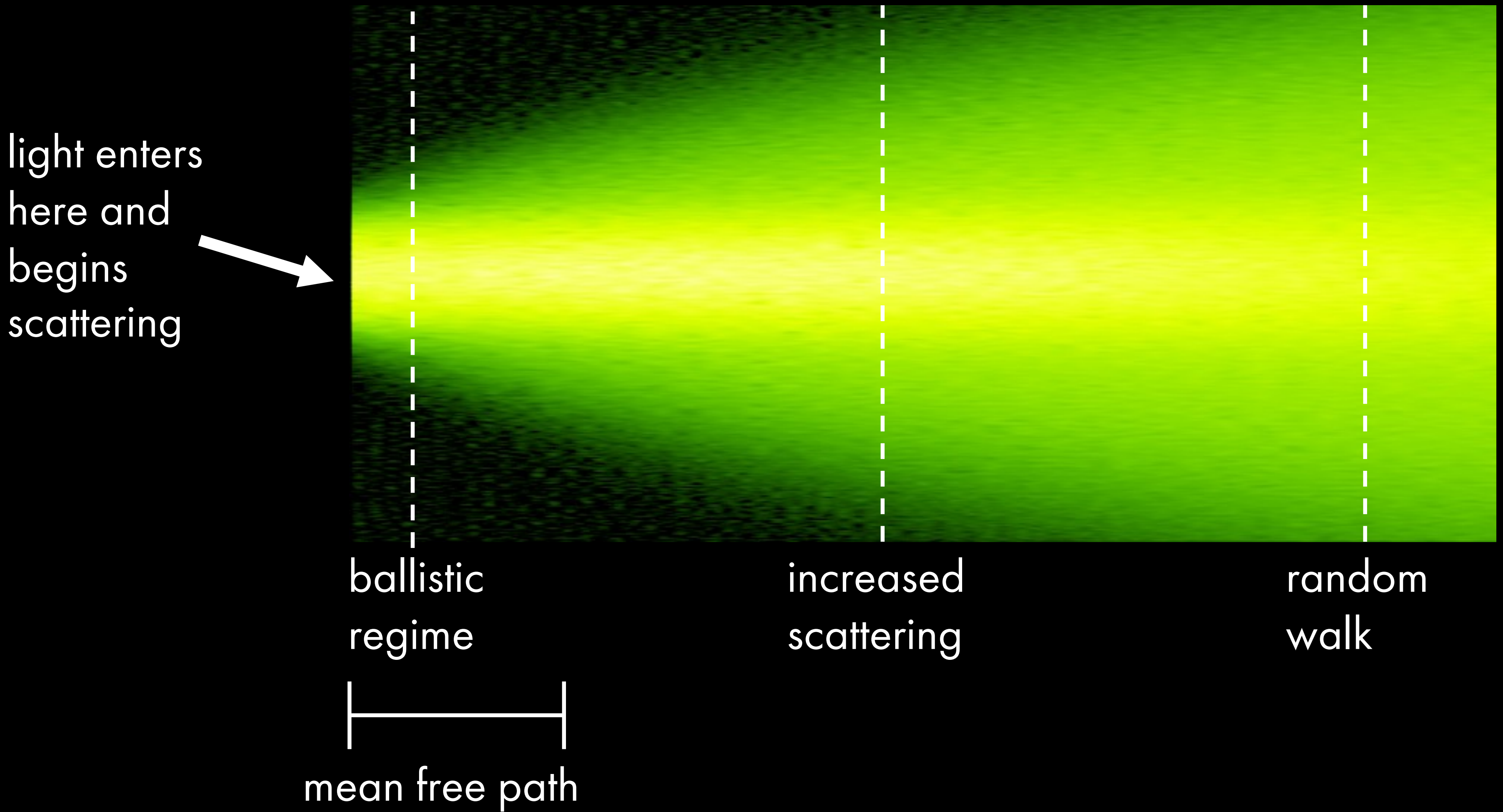
Challenges

- very few returning photons
- information is 'scrambled' by scattering

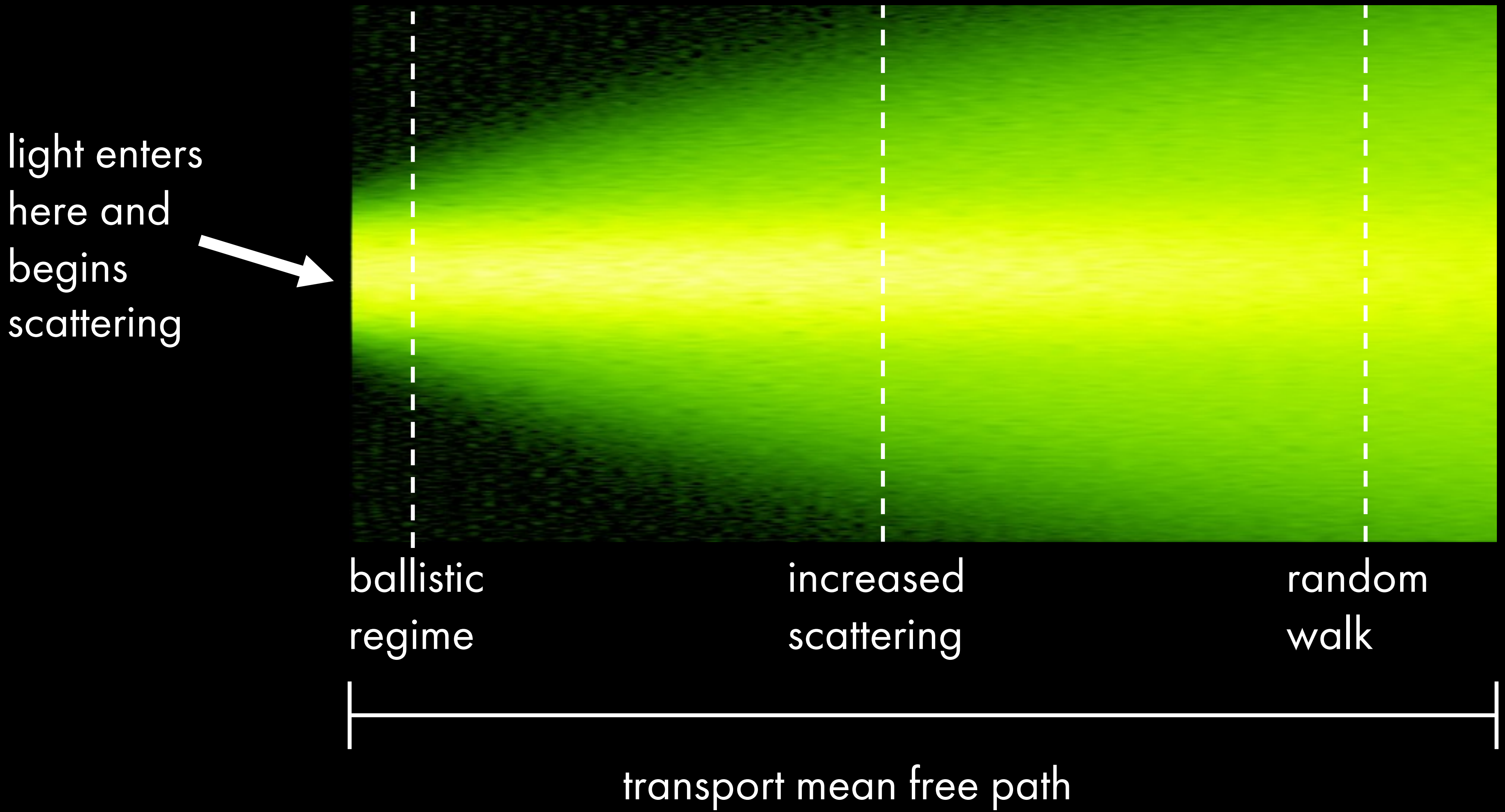
Imaging through scattering media



Imaging through scattering media



Imaging through scattering media



Imaging through scattering media

this work

> 6 transport mean free paths (TMFP)

Connection to Radiative Transfer

$$\underbrace{(\omega \cdot \nabla) L(\mathbf{x}, \omega)}_{\text{radiance in direction } \omega} = \underbrace{-\sigma_t(\mathbf{x}) L(\mathbf{x}, \omega)}_{\text{scattering/absorption}} + \underbrace{L_e(\mathbf{x}, \omega)}_{\text{emission}} + \underbrace{\sigma_s(\mathbf{x}) \int_{\mathcal{S}^2} f_p(\mathbf{x}, \omega, \omega') L(\mathbf{x}, \omega') \mathrm{d}\omega'}_{\text{in-scattering}}$$

Assumptions:

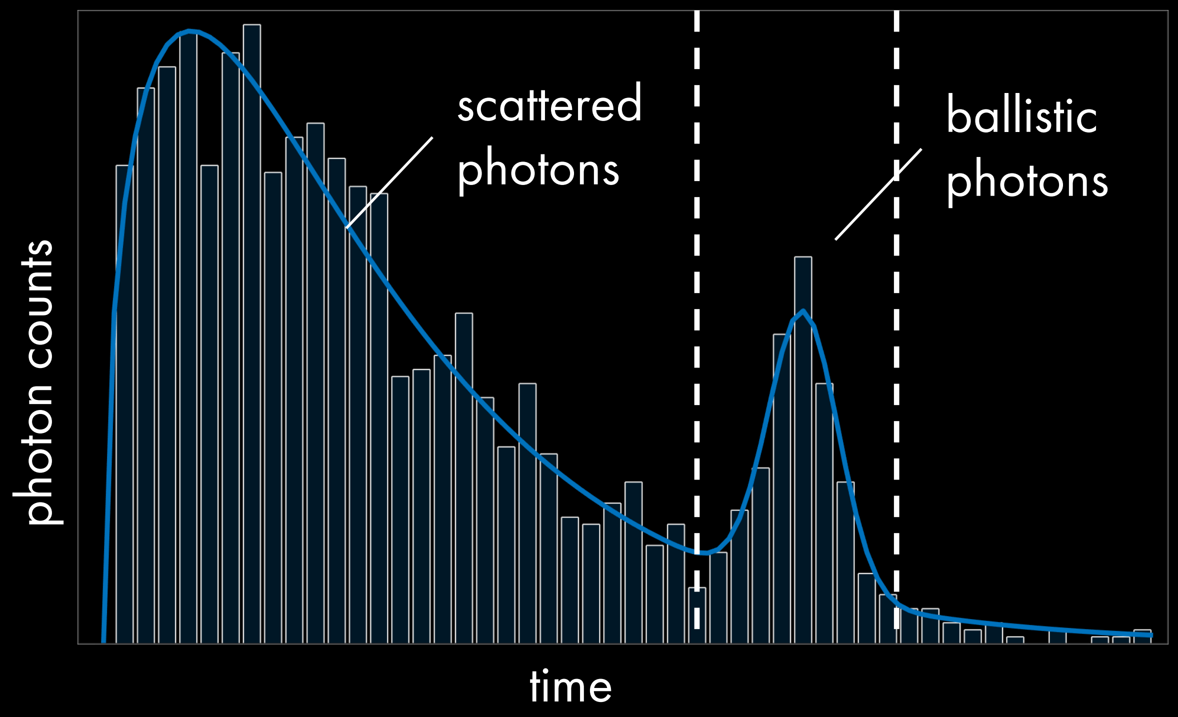
- low absorption
- highly-scattering regime

$$\frac{1}{c} \frac{\partial \Phi(\mathbf{x}, t)}{\partial t} + \sigma_a \Phi(\mathbf{x}, t) - \nabla \cdot [D \nabla \Phi(\mathbf{x}, t)] = L_e(\mathbf{x}, t)$$

diffusion equation

Related Work

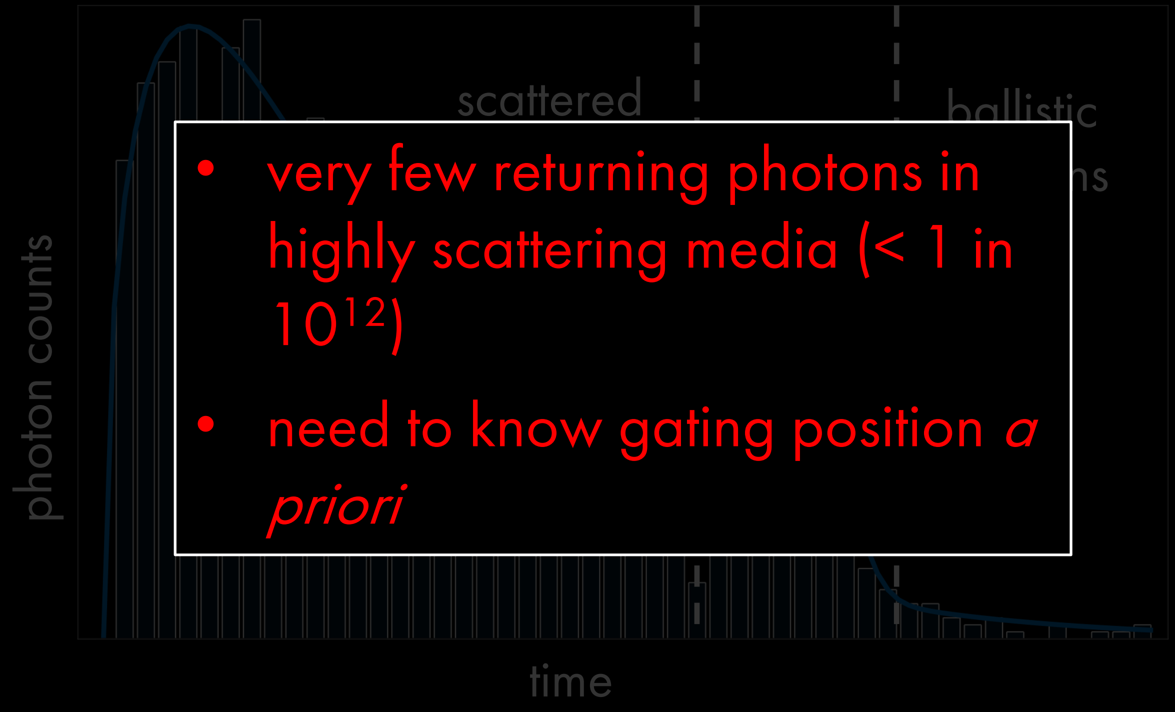
Ballistic imaging



[Wang '91], [Redo-Sanchez '16], [Satat '18], ...

Related Work

Ballistic imaging

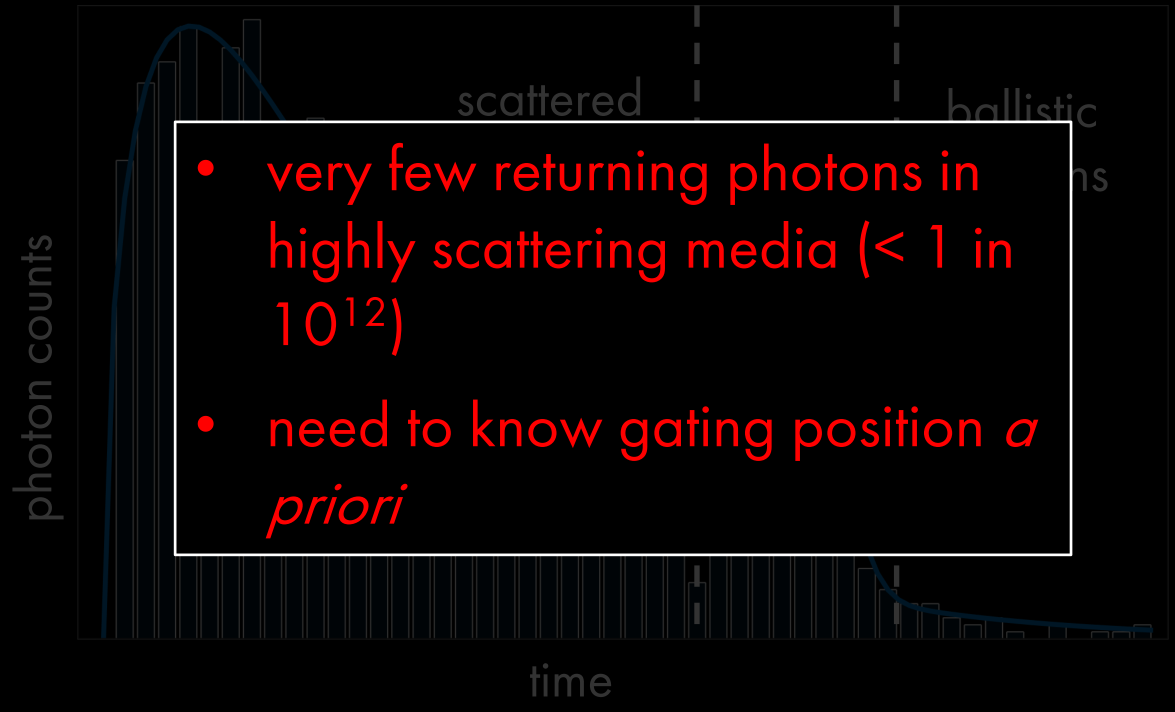


- very few returning photons in highly scattering media (< 1 in 10^{12})
- need to know gating position *a priori*

[Wang '91], [Redo-Sanchez '16], [Satat '18], ...

Related Work

Ballistic imaging

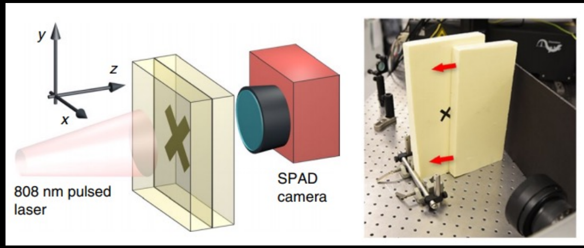


- very few returning photons in highly scattering media (< 1 in 10^{12})
- need to know gating position *a priori*

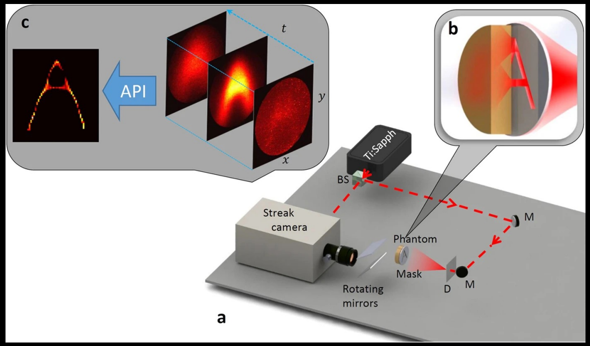
[Wang '91], [Redo-Sanchez '16], [Satat '18], ...

Diffuse Optical Tomography

2D

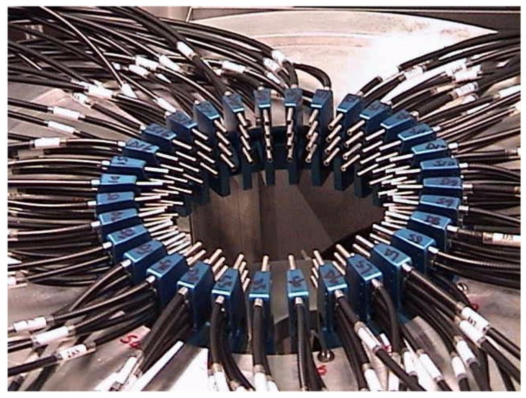


[Lyons '19]



[Satat '16]

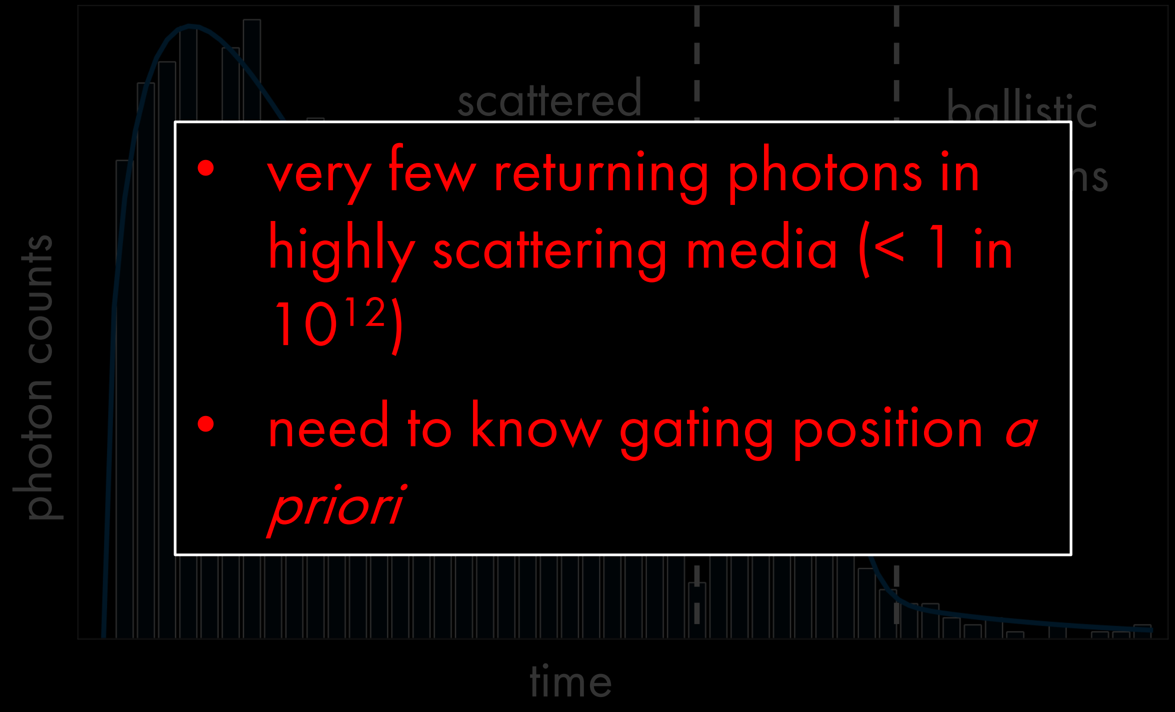
3D



[Hajihashemi '12]

Related Work

Ballistic imaging

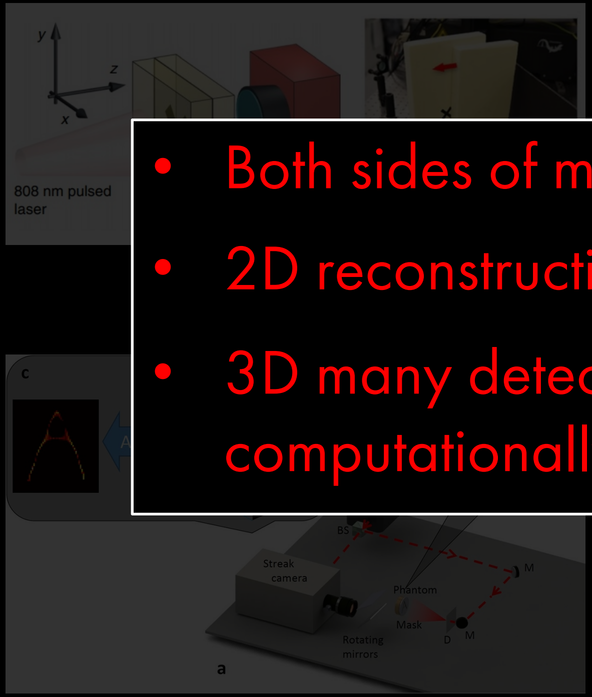


- very few returning photons in highly scattering media (< 1 in 10^{12})
- need to know gating position *a priori*

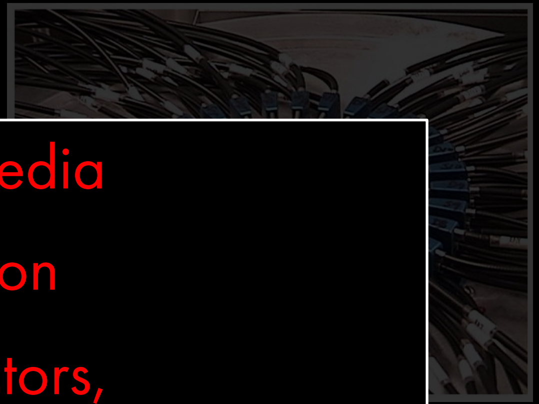
[Wang '91], [Redo-Sanchez '16], [Satat '18], ...

Diffuse Optical Tomography

2D



3D

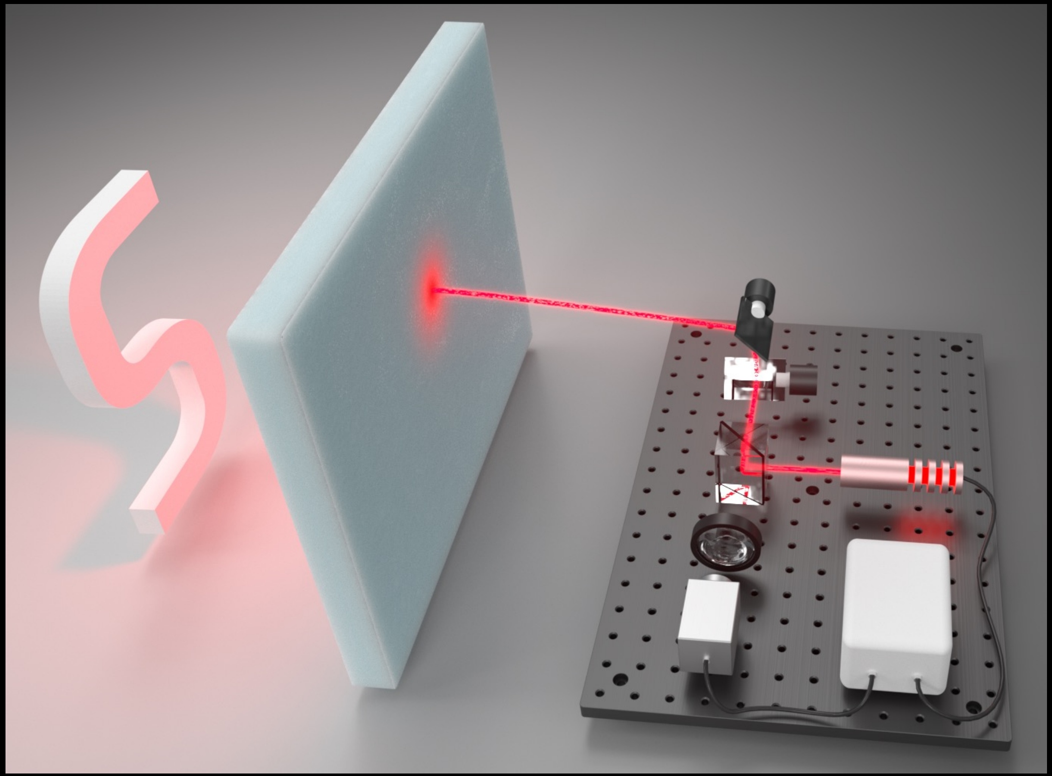


- Both sides of media
- 2D reconstruction
- 3D many detectors, computationally expensive

[Satat '16]

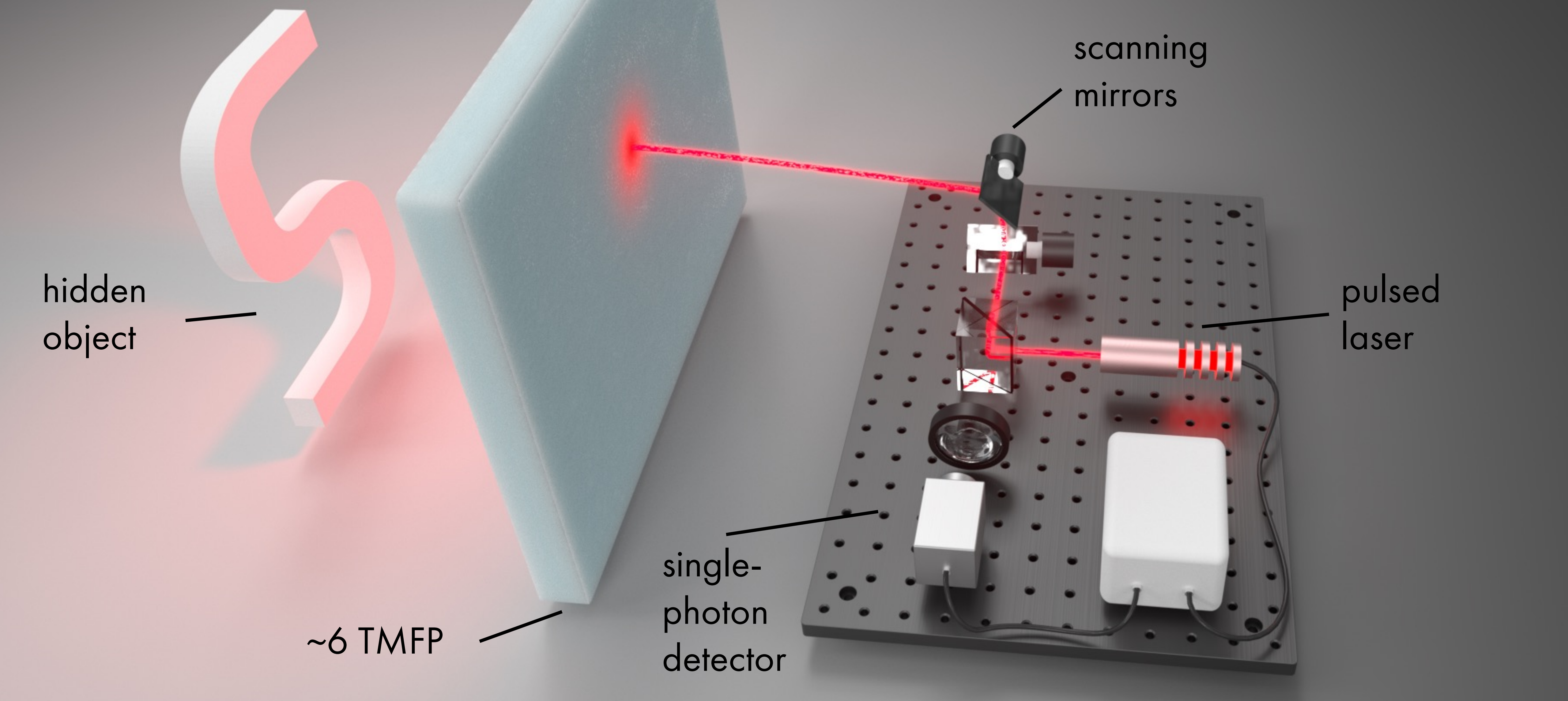
This work

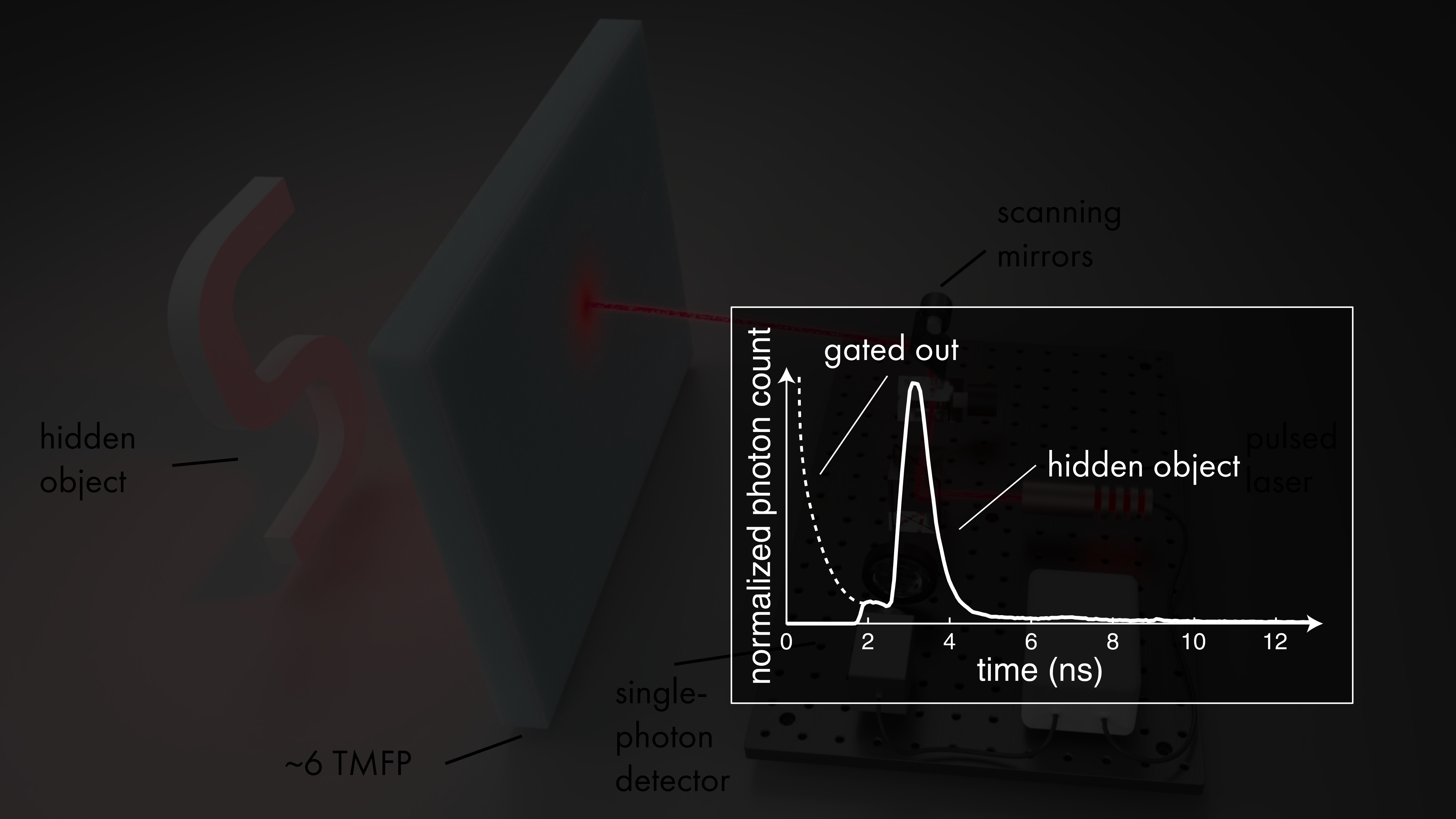
- Invert light transport
- non-invasive, reflection mode
- efficient 3D reconstruction at meter scales without *a priori* knowledge of target



Assumptions:

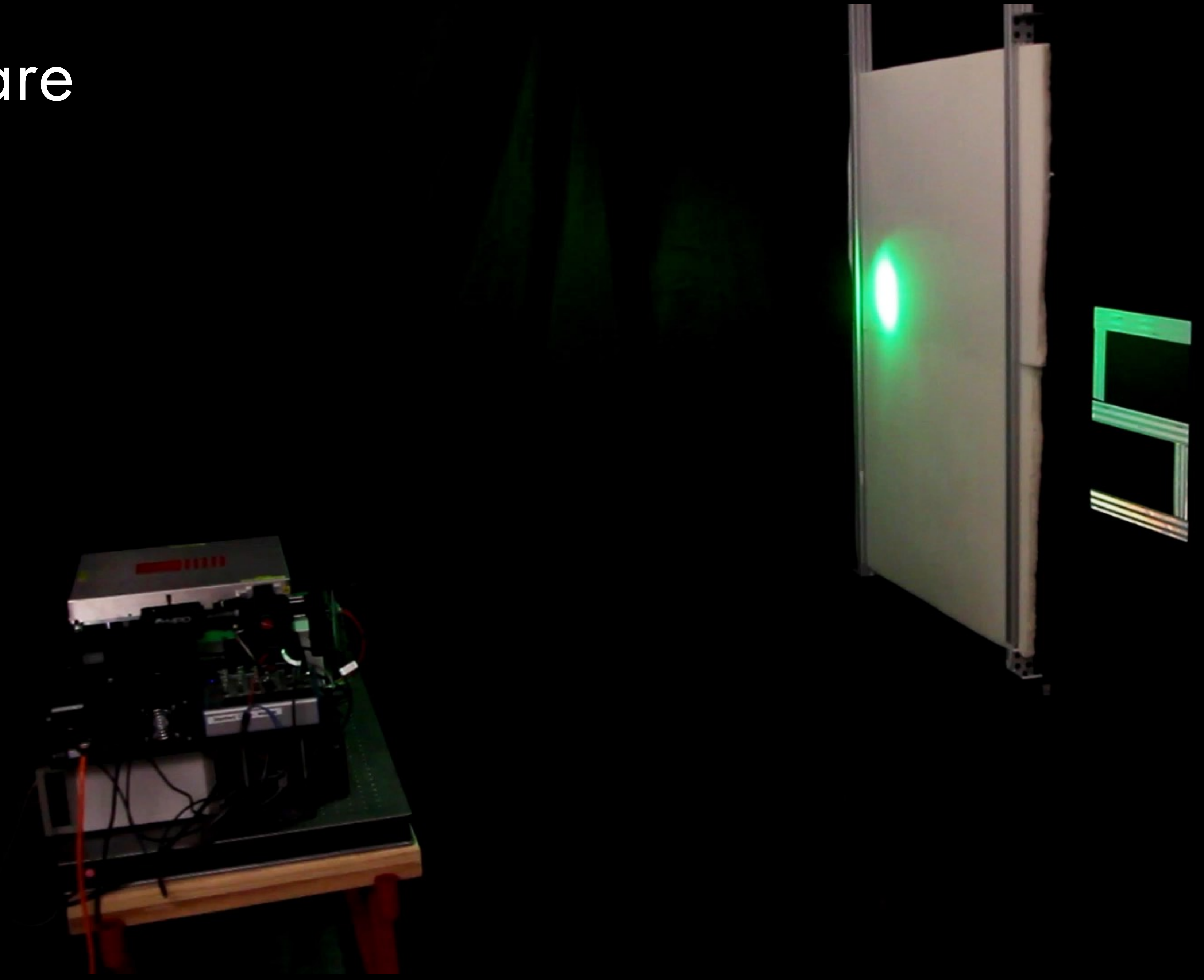
- object at a distance behind scattering media
- scattering media is static



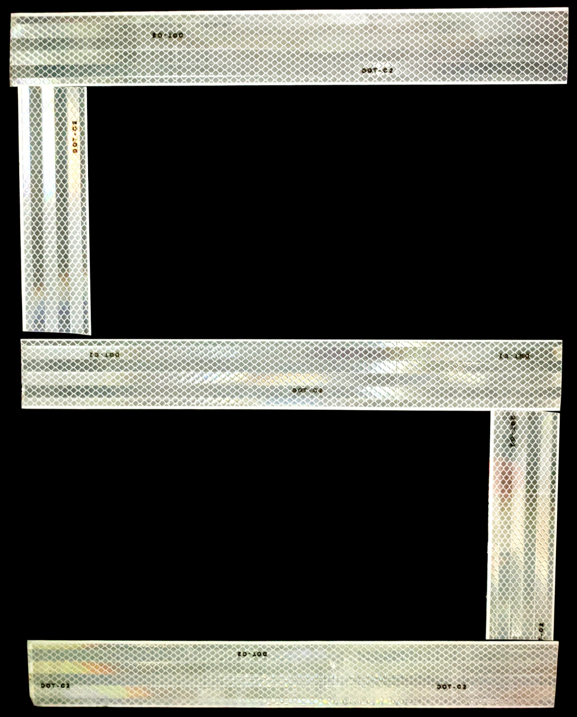


hidden object
~6 TMFP
single-photon detector
scanning mirrors
pulsed laser
gated out
hidden object
normalized photon count
time (ns)
0 2 4 6 8 10 12

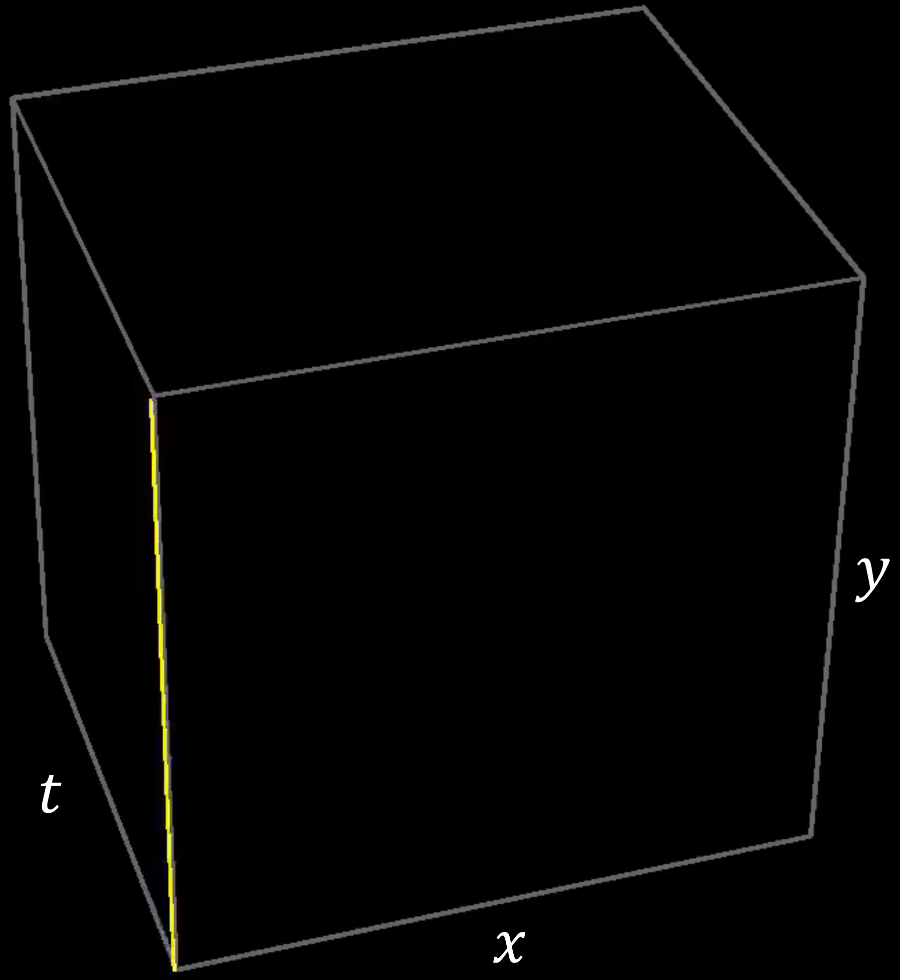
Hardware



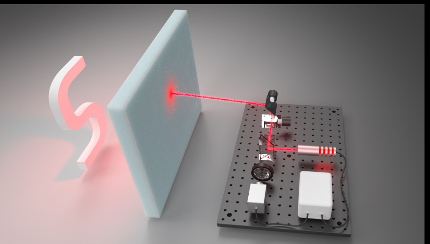
Results



hidden object

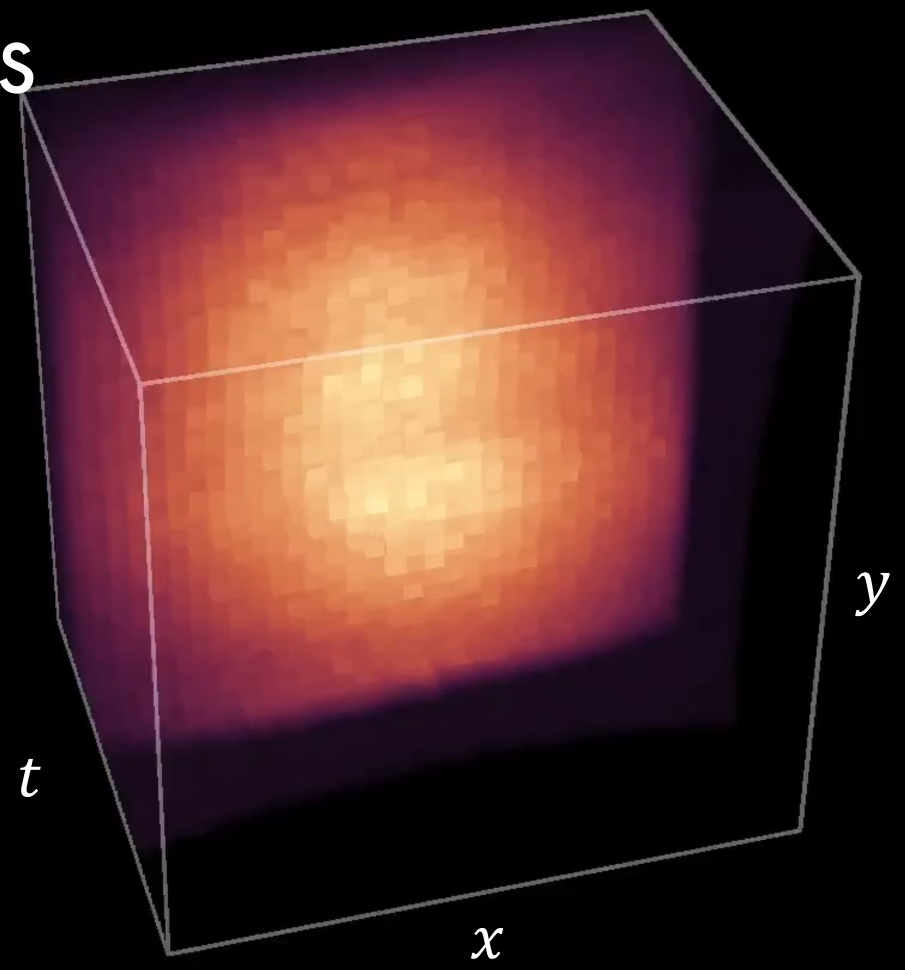


captured measurements
(0.6 m × 0.6 m × 3.3 ns)

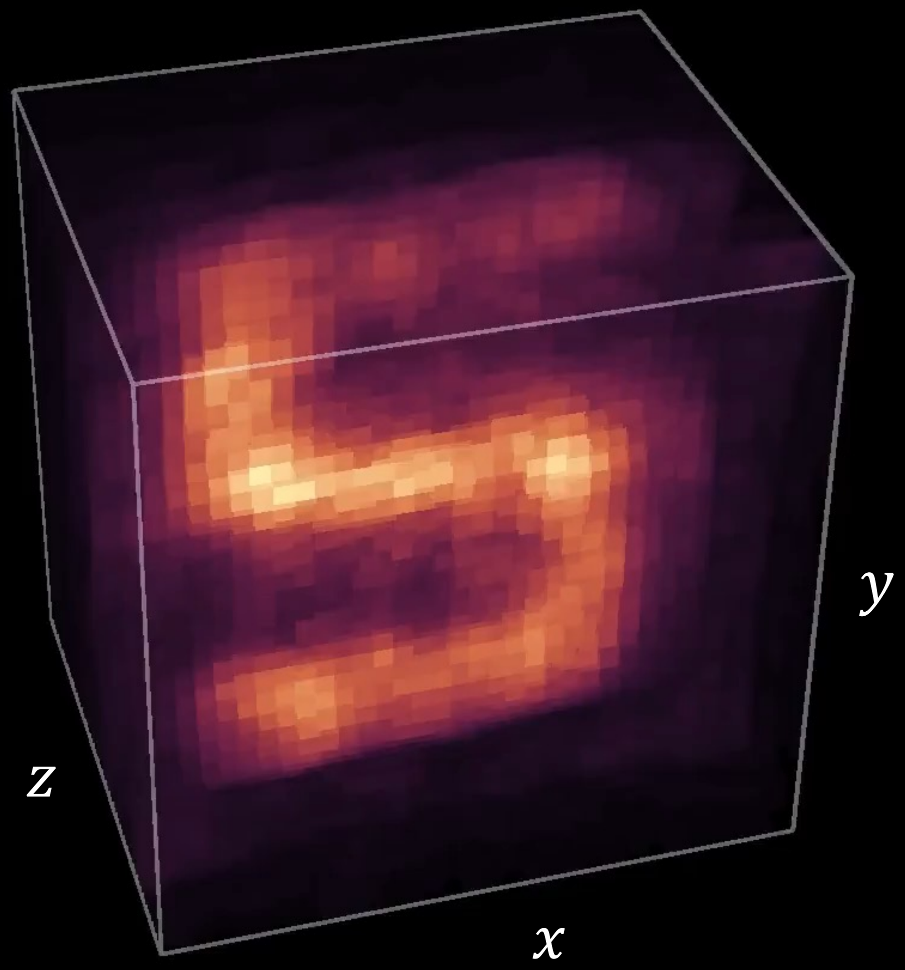


total acquisition time: 1 min. (60 ms/sample)

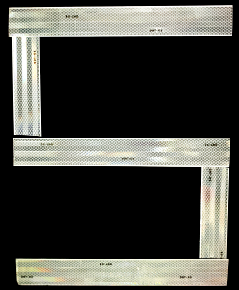
Results



captured measurements

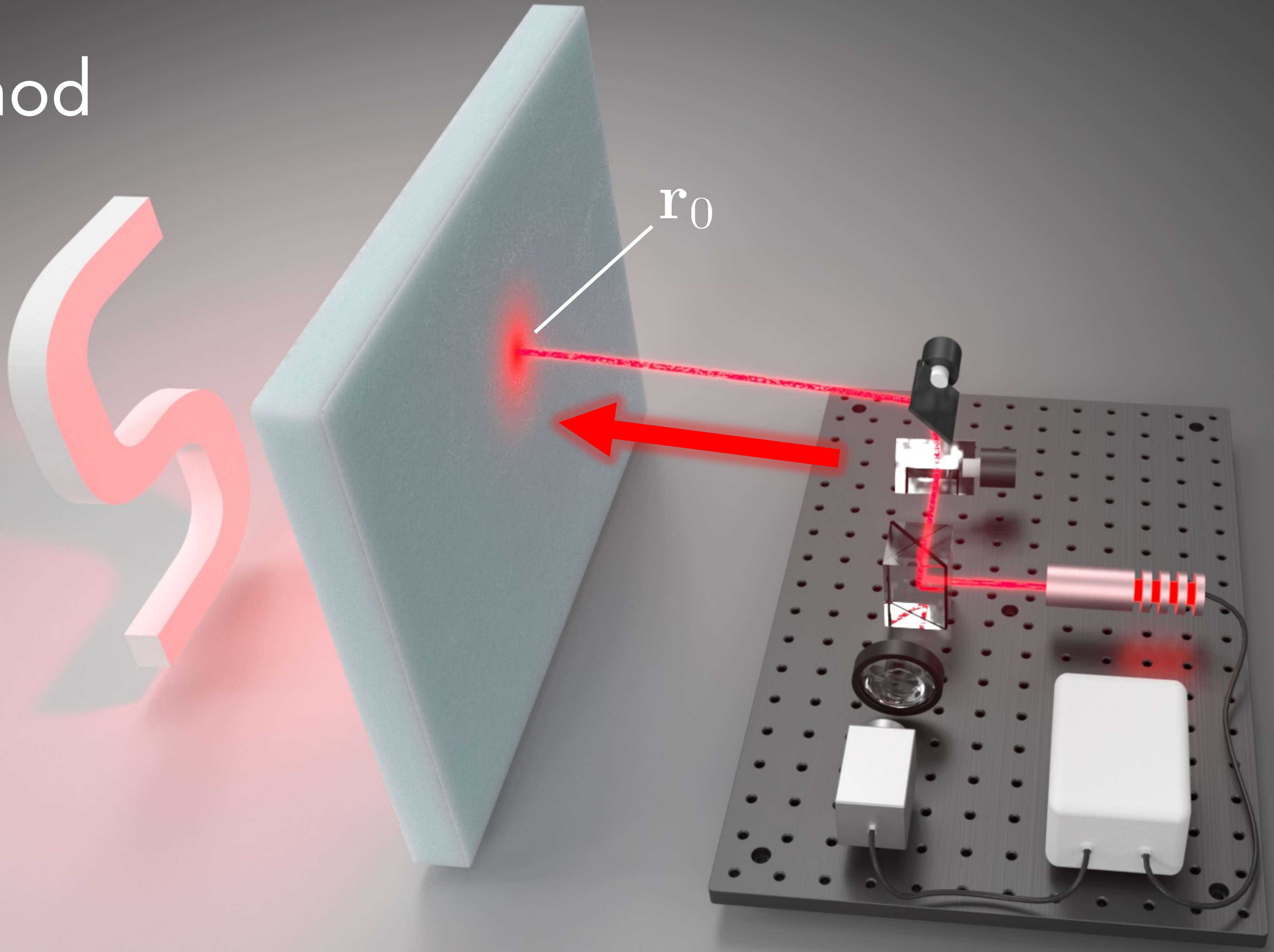


reconstruction (50 ms)
(0.6 m × 0.6 m × 0.5 m)

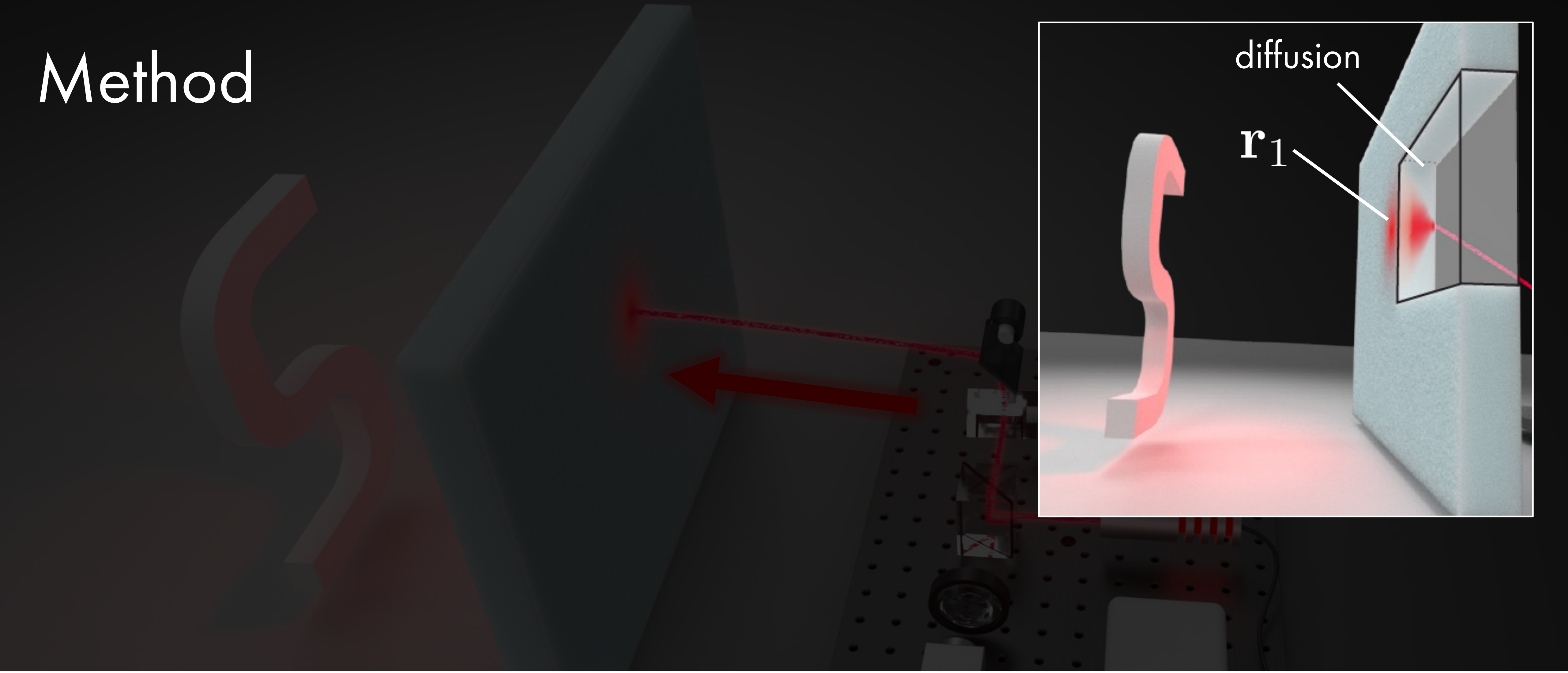


total acquisition time: 1 min. (60 ms/sample)

Method

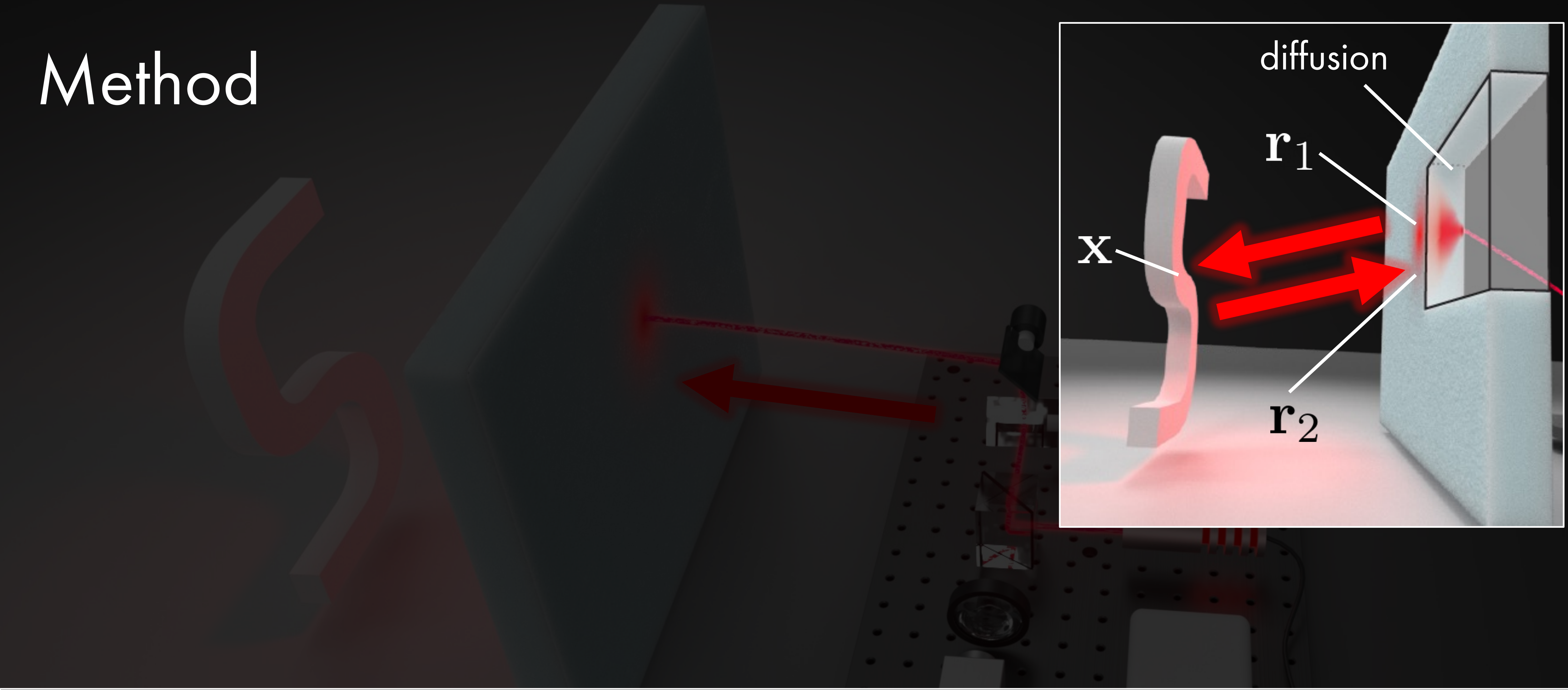


Method



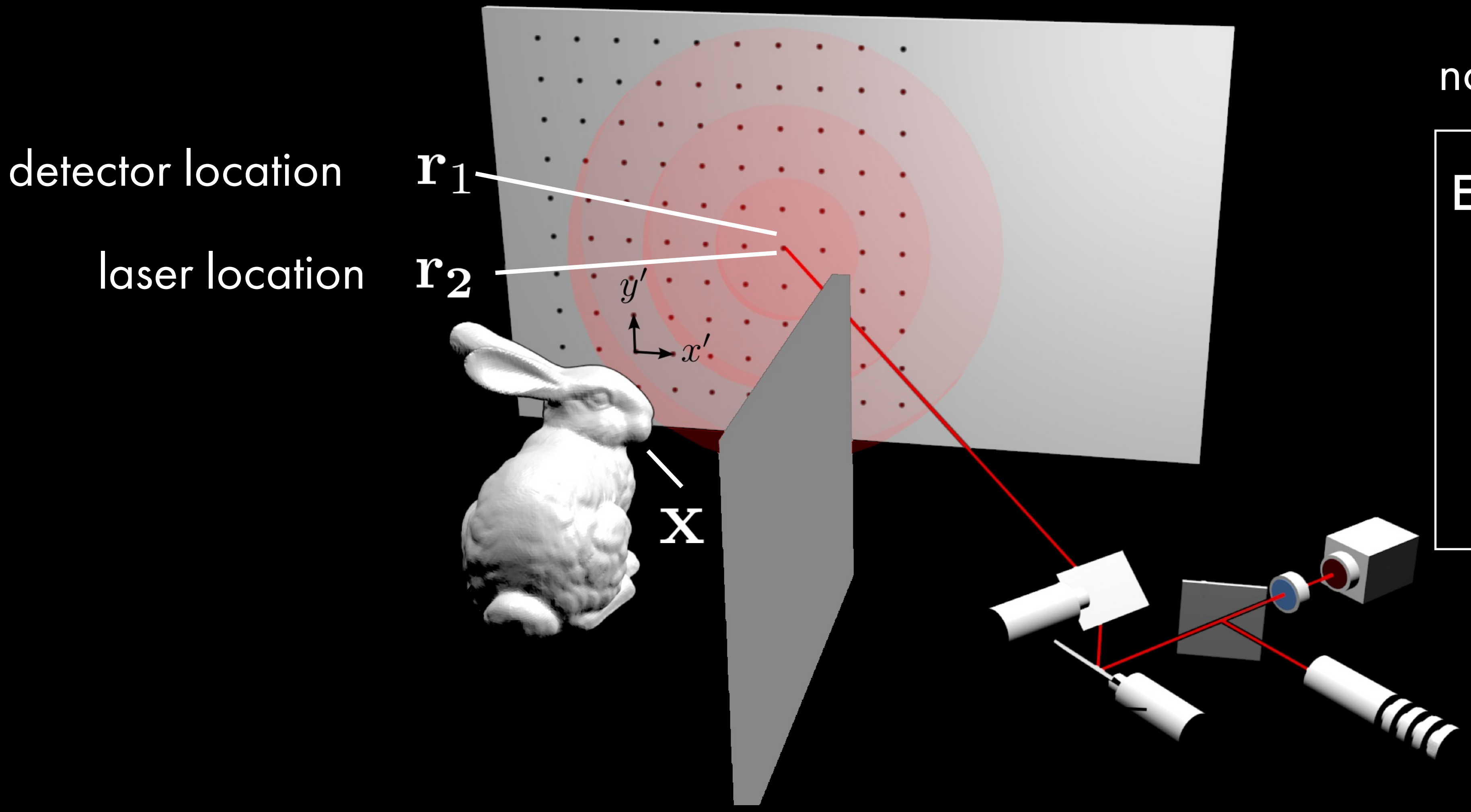
$$\phi(t, \mathbf{r_0}, \mathbf{r_1}) = \frac{c}{(4\pi Dct)^{3/2}} \exp\left(-\frac{\|\mathbf{r_1} - \mathbf{r_0}\|_2^2}{4Dct} - \mu_a ct\right)$$

Method



How to efficiently model free space propagation?

Method

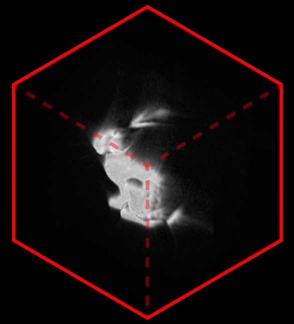


non-line-of-sight imaging

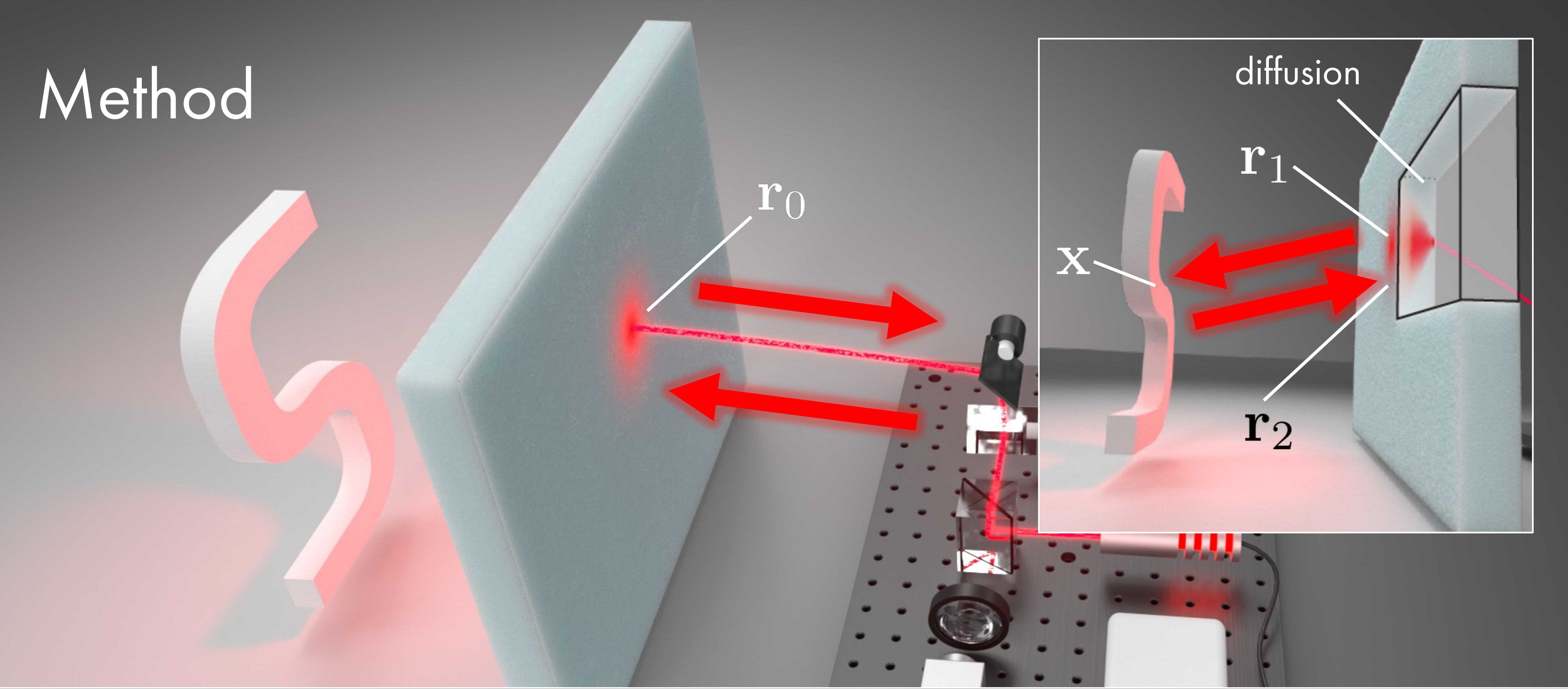
Efficient inversion if:

$$\mathbf{r_1} = \mathbf{r_2}$$

$$I(t, \mathbf{r_1}, \mathbf{r_1})$$

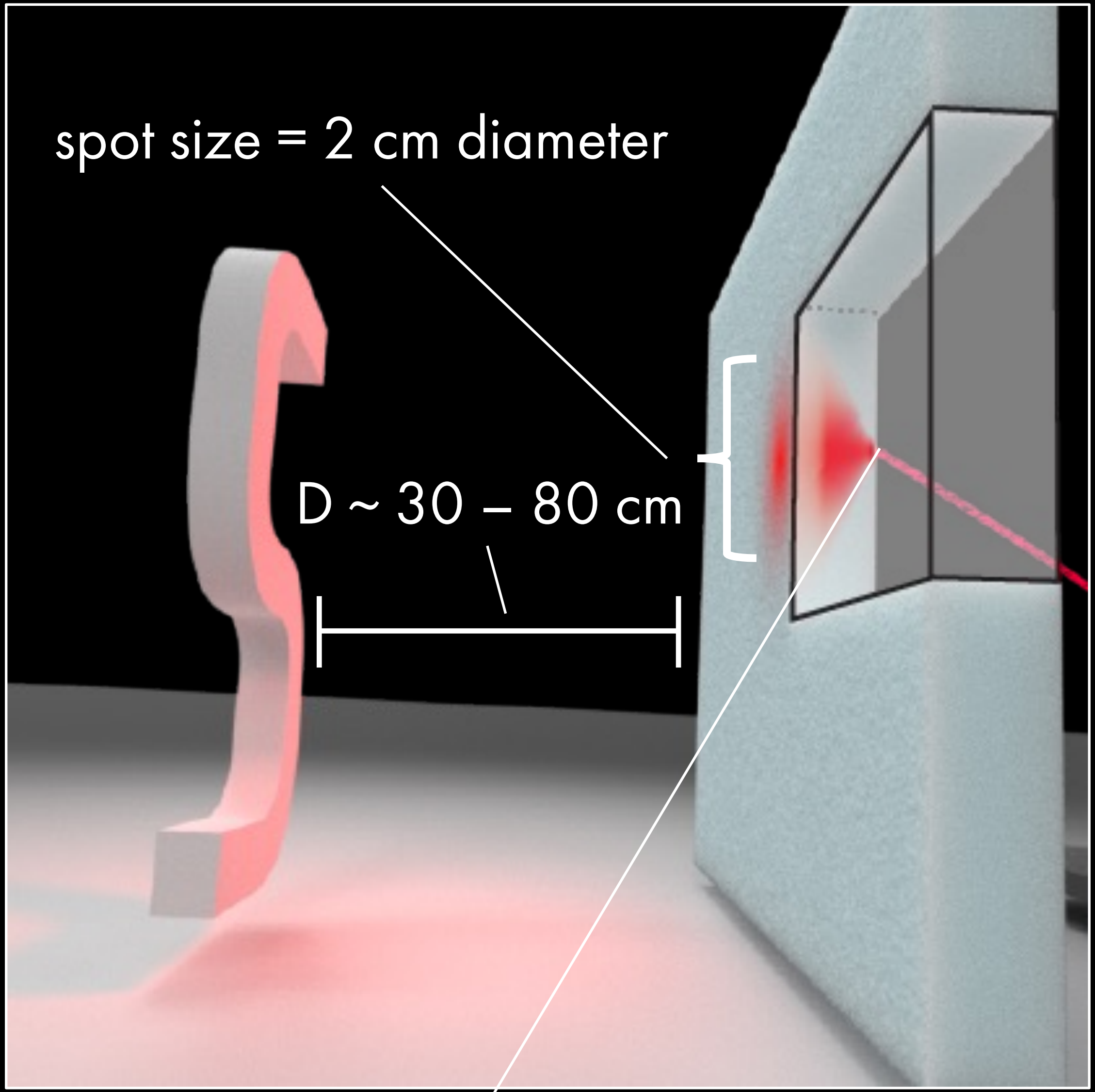


Method



$$\phi(t, \mathbf{r_0}, \mathbf{r_1}) = \frac{c}{(4\pi Dct)^{3/2}} \exp\left(-\frac{\|\mathbf{r_1} - \mathbf{r_0}\|_2^2}{4Dct} - \mu_a ct\right)$$

Method

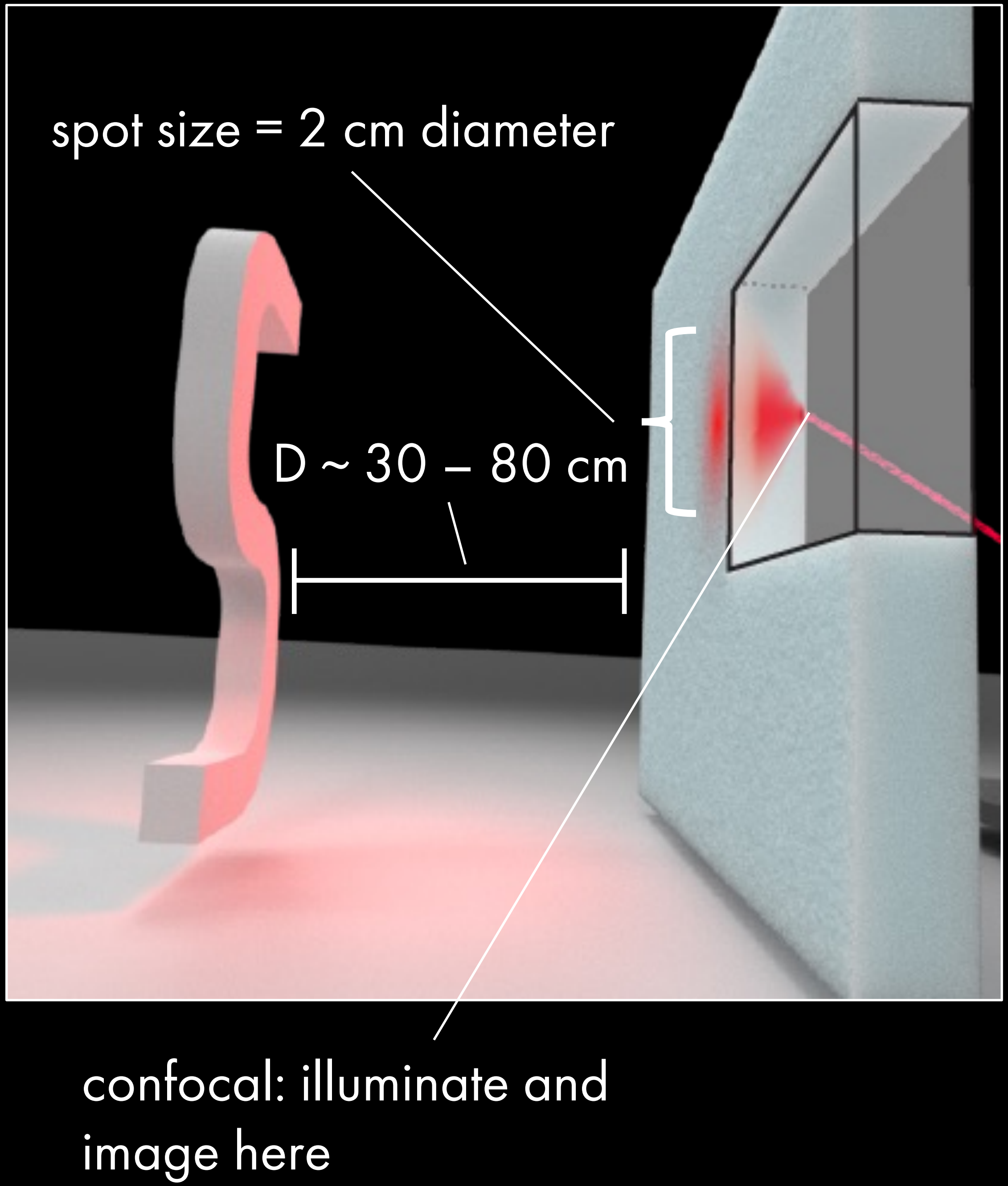

spot size = 2 cm diameter
D ~ 30 – 80 cm
confocal: illuminate and image here

Method

Approximation:

Approximate measured light as scattering back to the same spot. $\mathbf{r_1} = \mathbf{r_2}$

Error ~ (spot size)^2 / (2 * distance) << 1 cm



Method

Approximation:

Approximate measured light as scattering back to the same spot. $\mathbf{r_1} = \mathbf{r_2}$

Error ~ (spot size)^2 / (2 * distance) << 1 cm

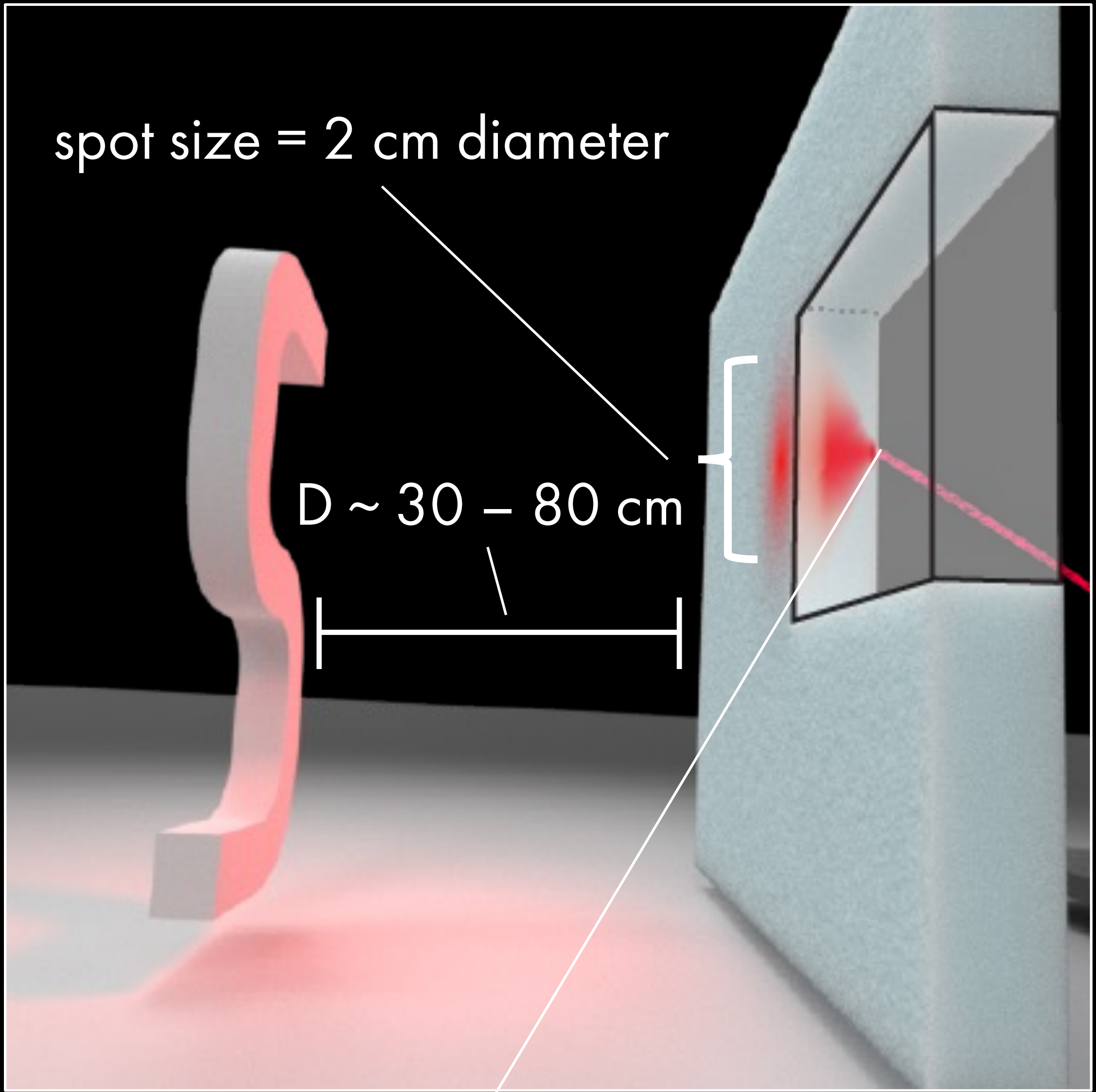
measurements

$$\hat{\tau}(t, \mathbf{r_0}) =$$

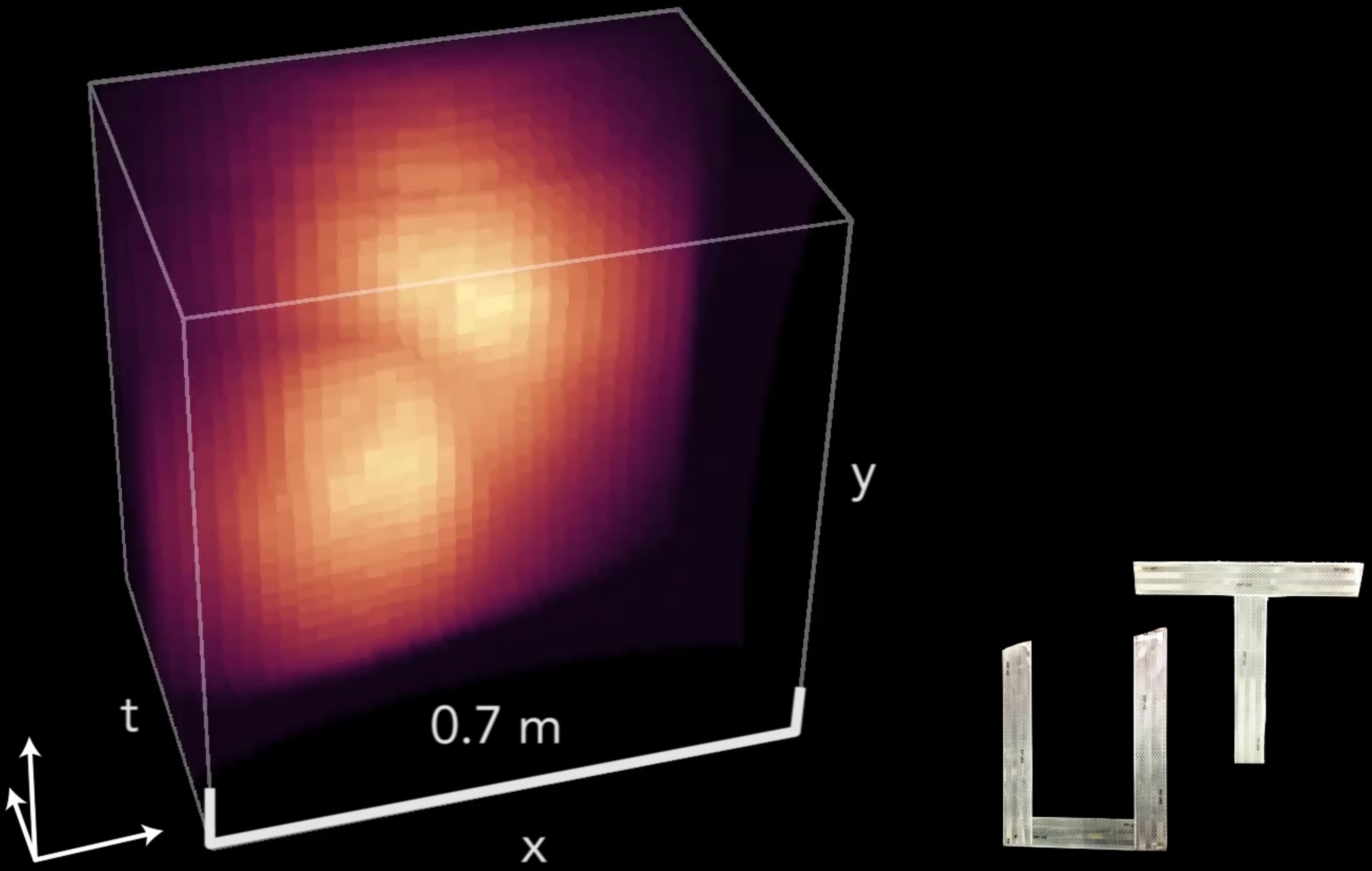
$$\phi(t, \mathbf{r_0}, \mathbf{r_1}) * \phi(t, \mathbf{r_0}, \mathbf{r_1}) * I(t, \mathbf{r_1}, \mathbf{r_1})$$

diffusion kernels NLOS model

Can use efficient NLOS inversion!



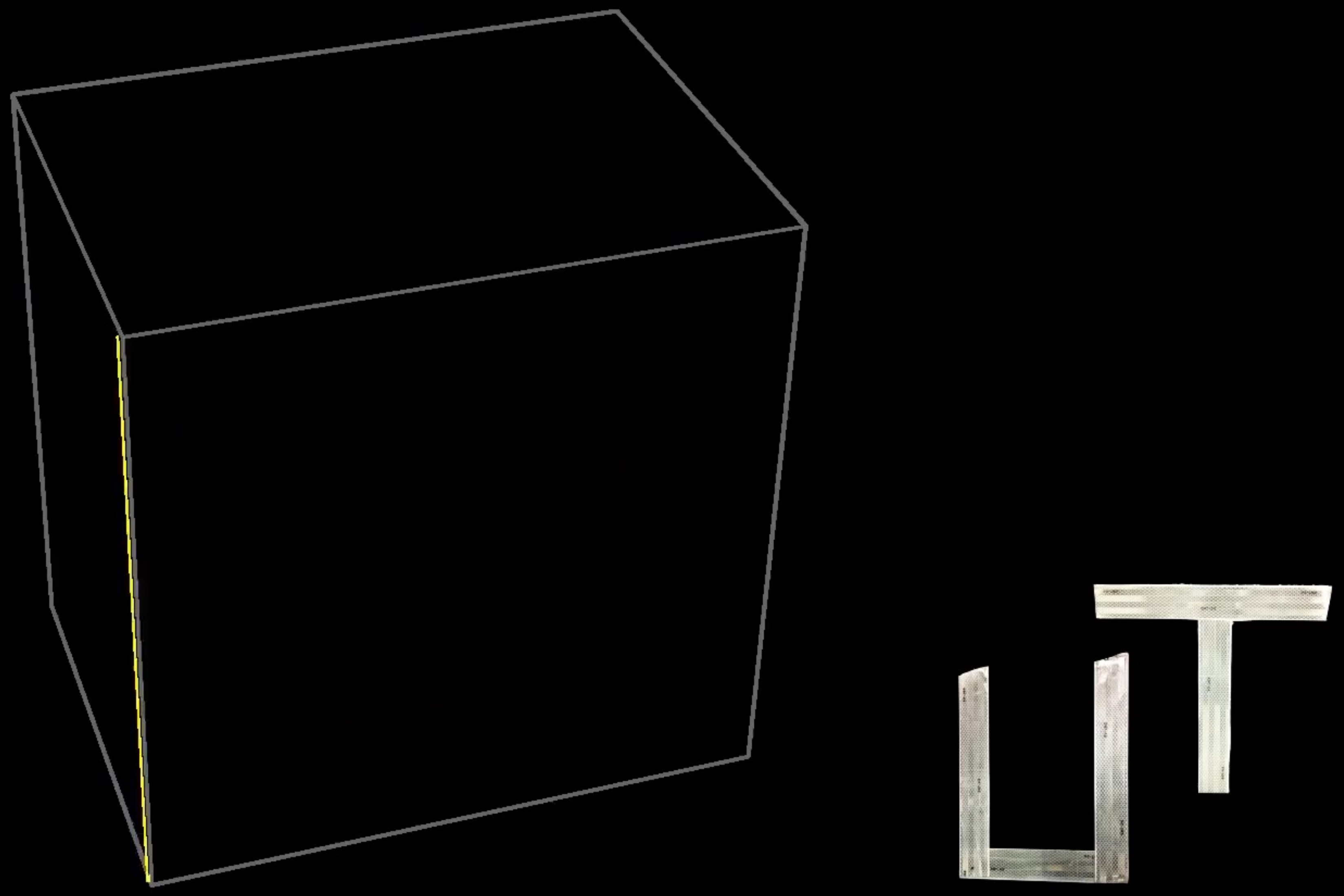
Results



measurements

Lindell et al., Nat. Commun. 2020

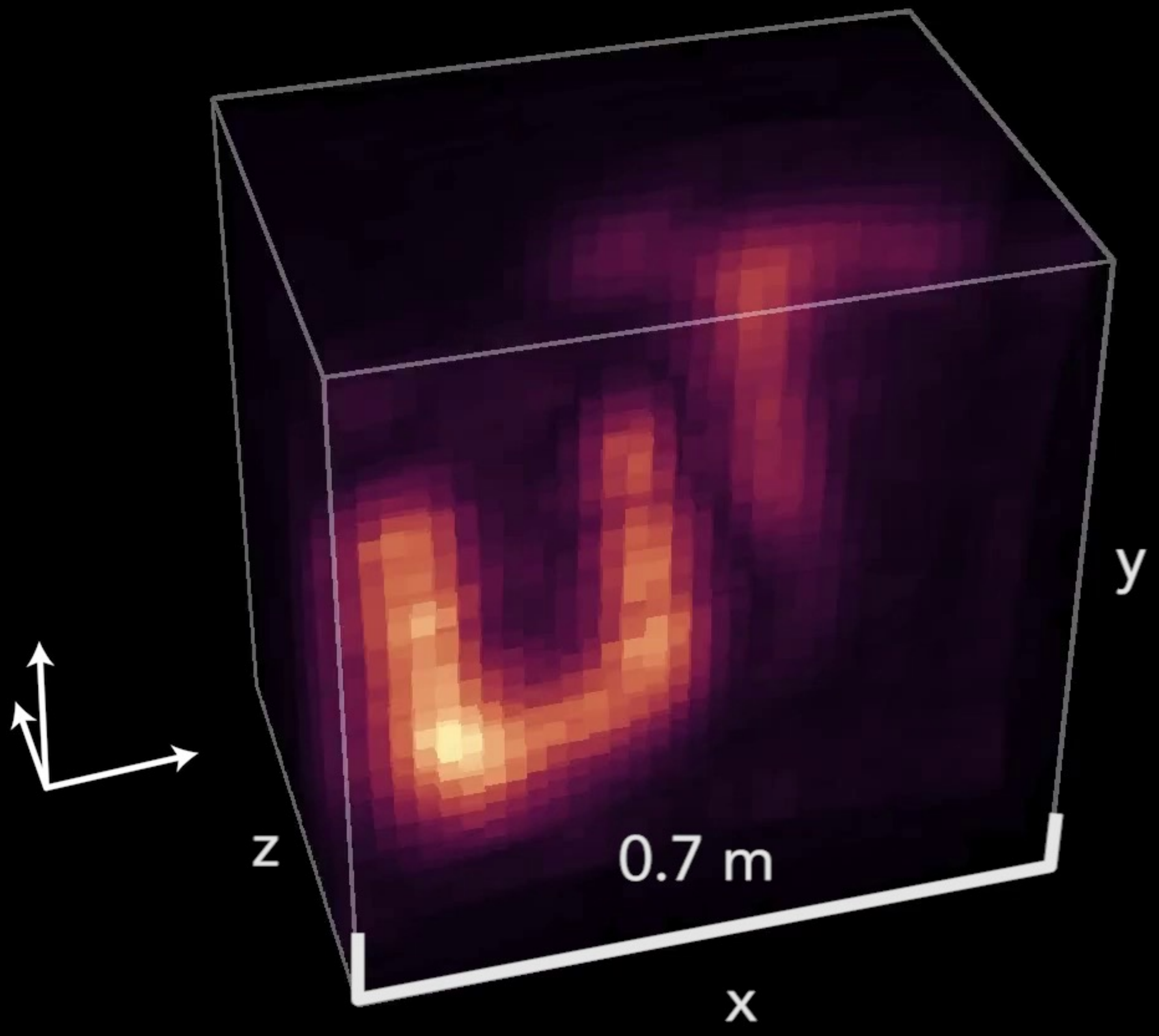
Results

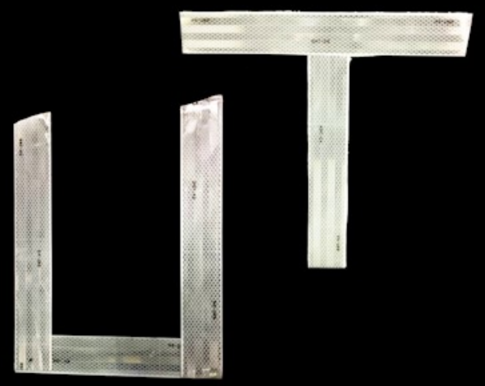


measurements

Lindell et al., Nat. Commun. 2020

Results

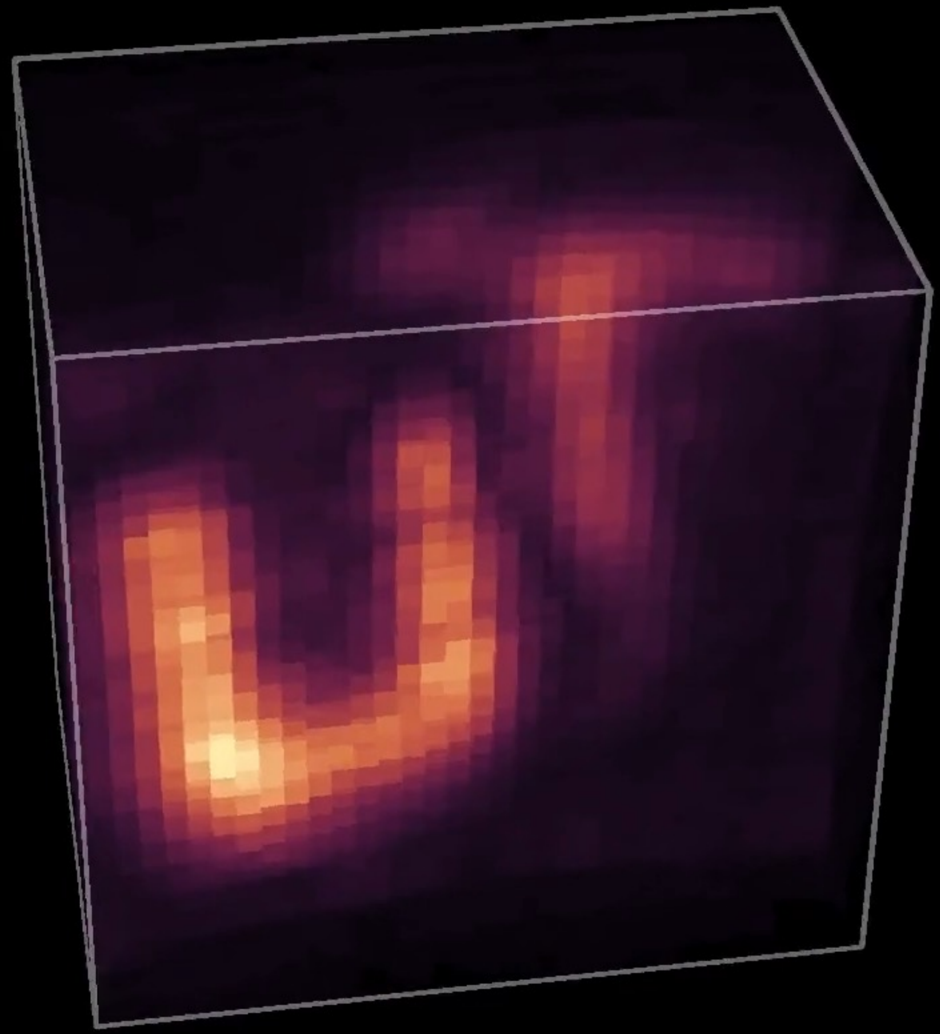


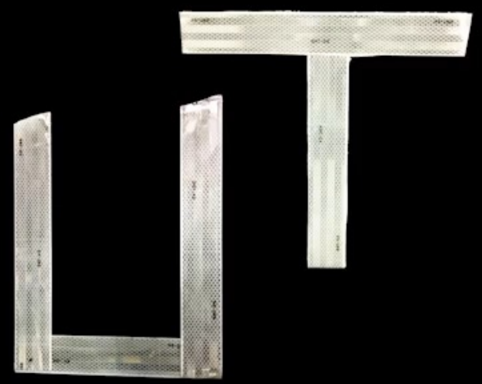


reconstruction

Lindell et al., Nat. Commun. 2020

Results



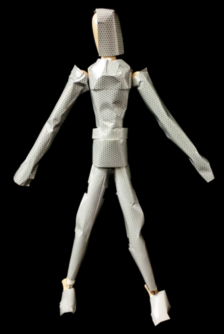


reconstruction

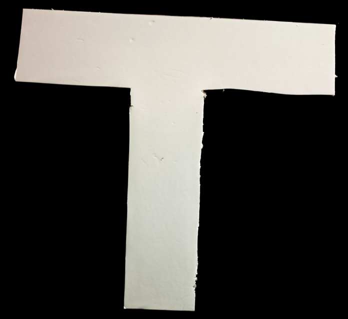
Lindell et al., Nat. Commun. 2020

Results

traffic cones



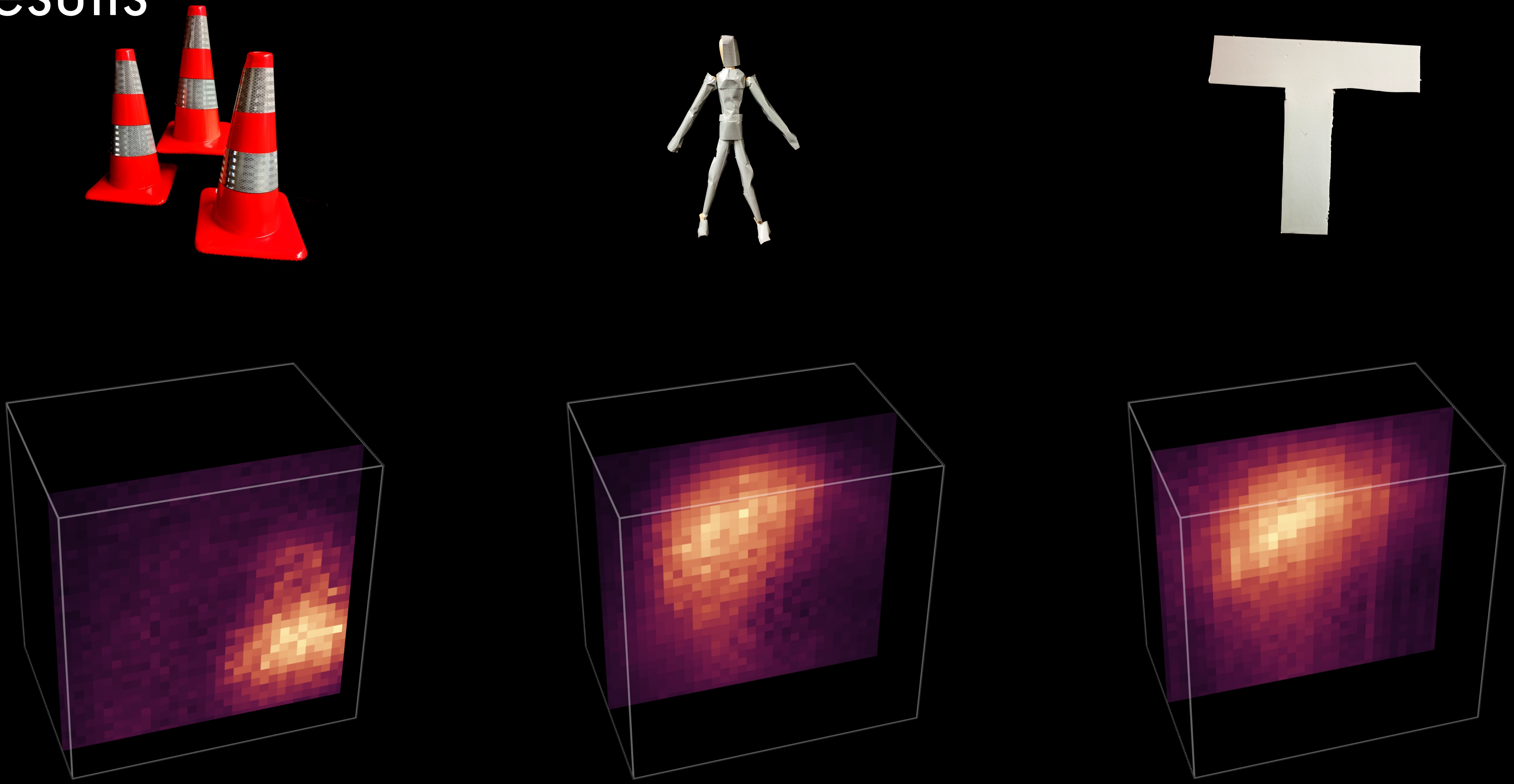
reflective mannequin



diffuse letter

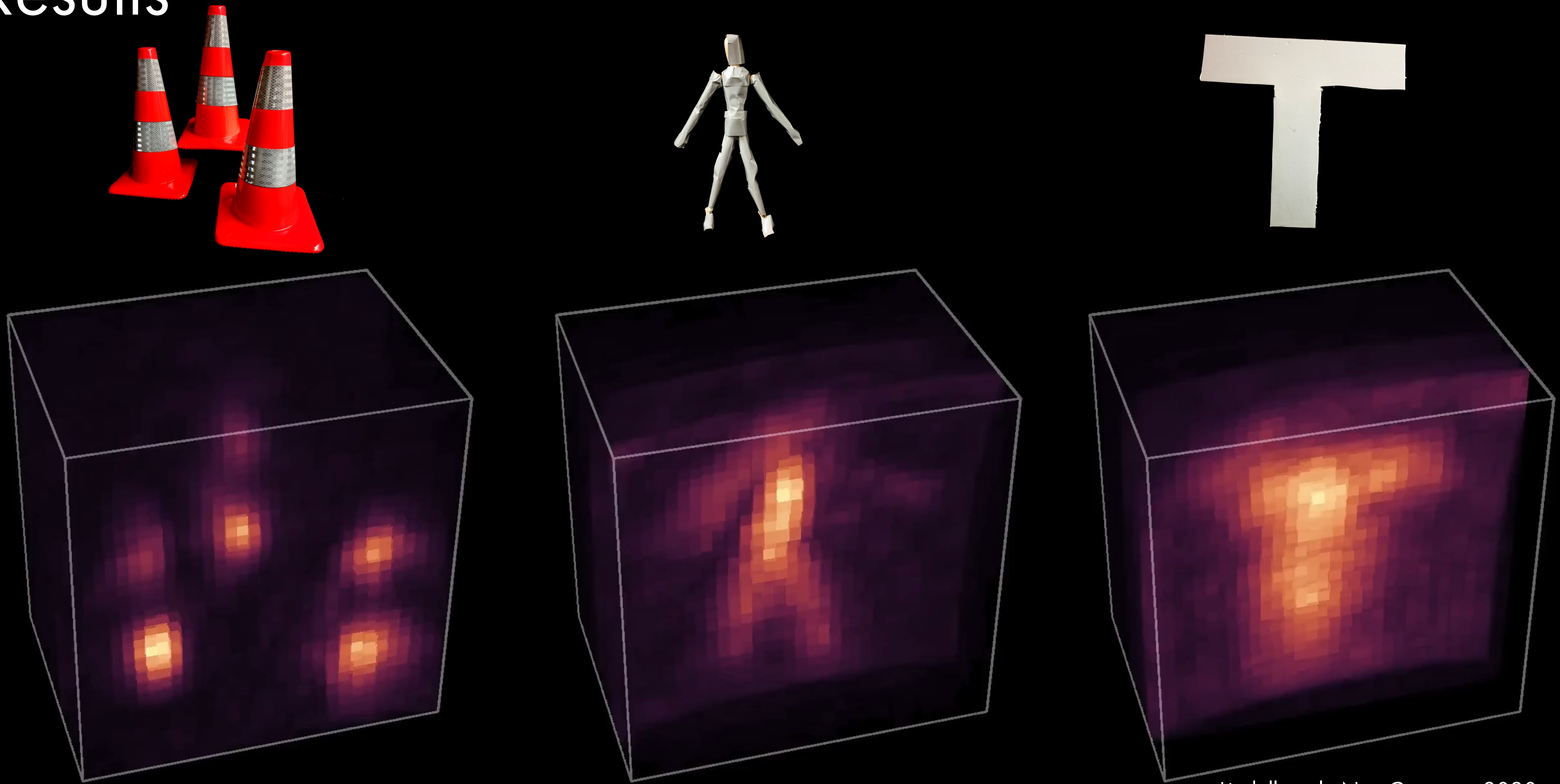
Lindell et al., Nat. Commun. 2020

Results



Lindell et al., Nat. Commun. 2020

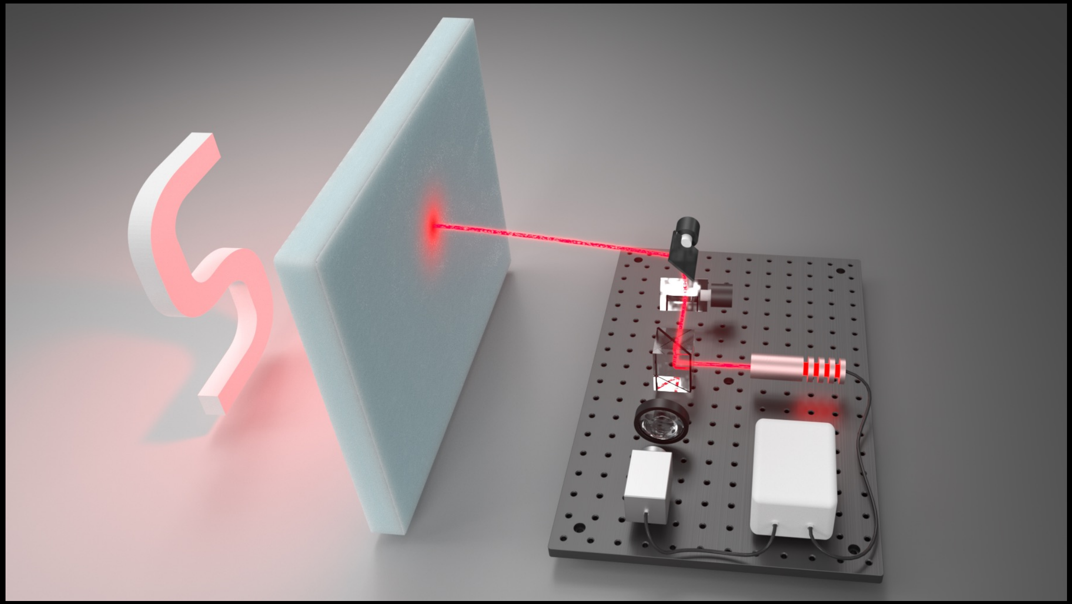
Results



Lindell et al., Nat. Commun. 2020

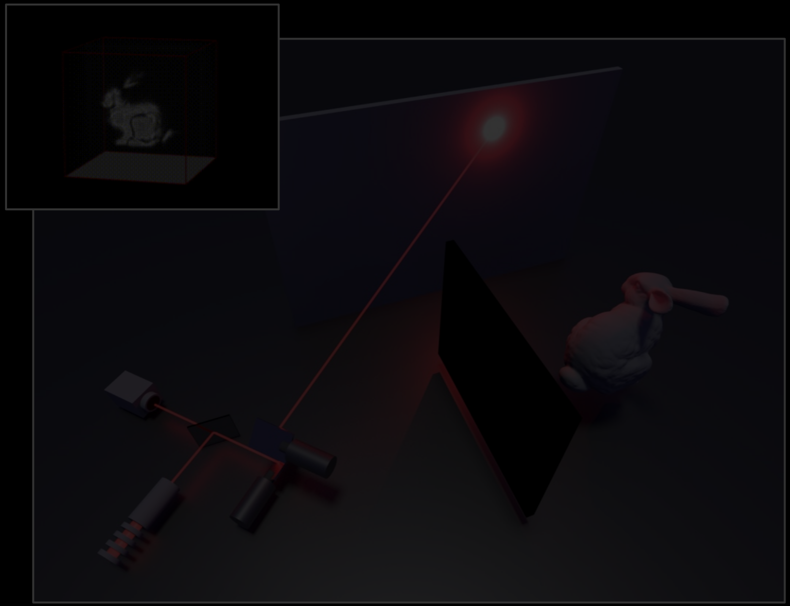
Outlook

- efficient method for 3D imaging through scattering media based on DOT
- works without *a priori* knowledge of target position
- What's next?
    - embedded, dynamic media
    - priors, AI algorithms



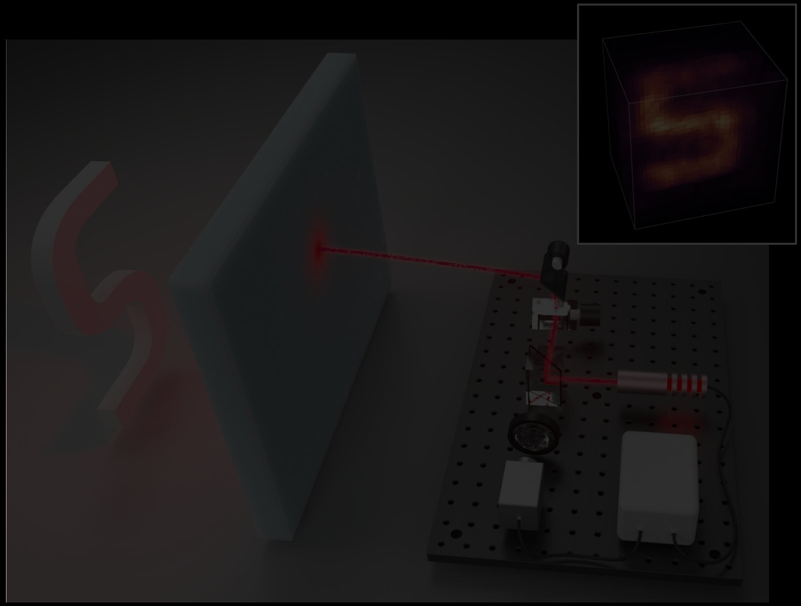
Overview

Non-Line-of-Sight Imaging



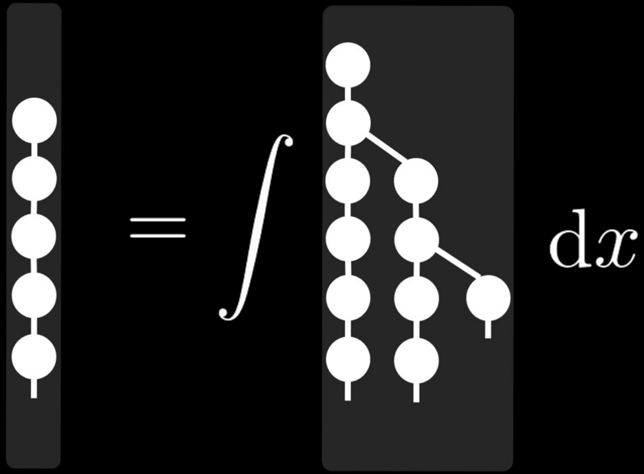
Nature '18
SIGGRAPH '19
CVPR '19
ACM Trans. Graph. '20
CVPR '20
IEEE TCI '21

Imaging through Scattering Media

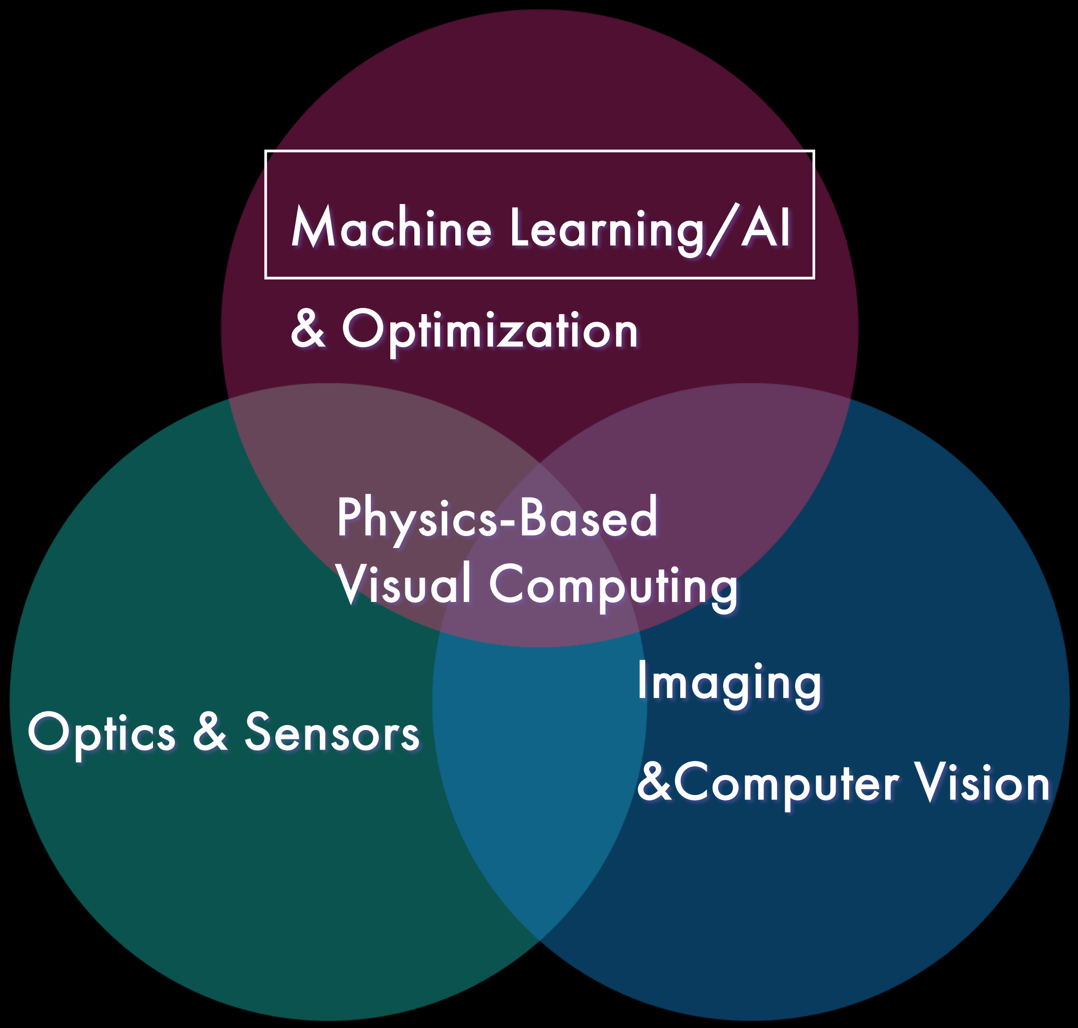


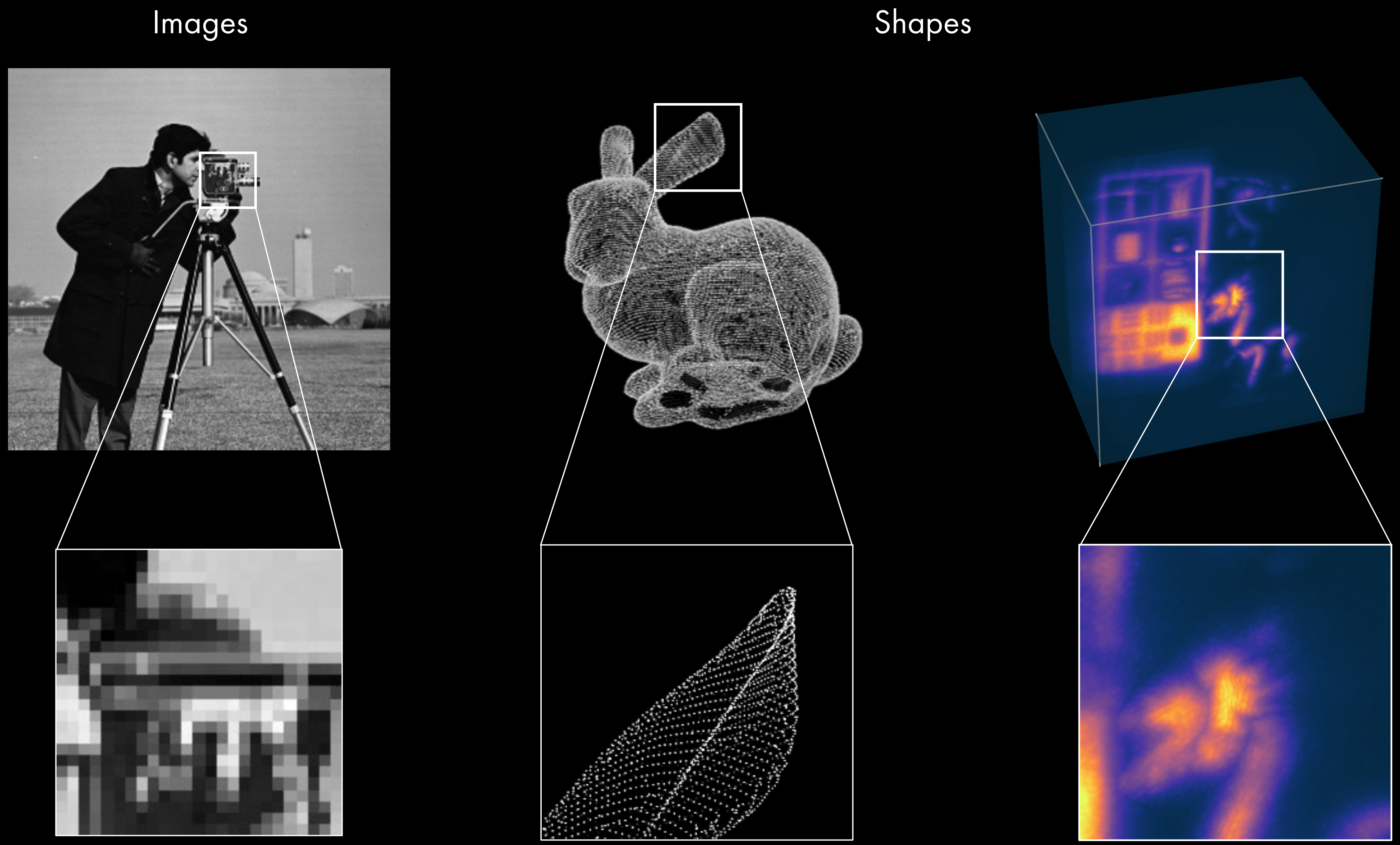
Nature Communications '20

Physics-based AI & Neural Rendering

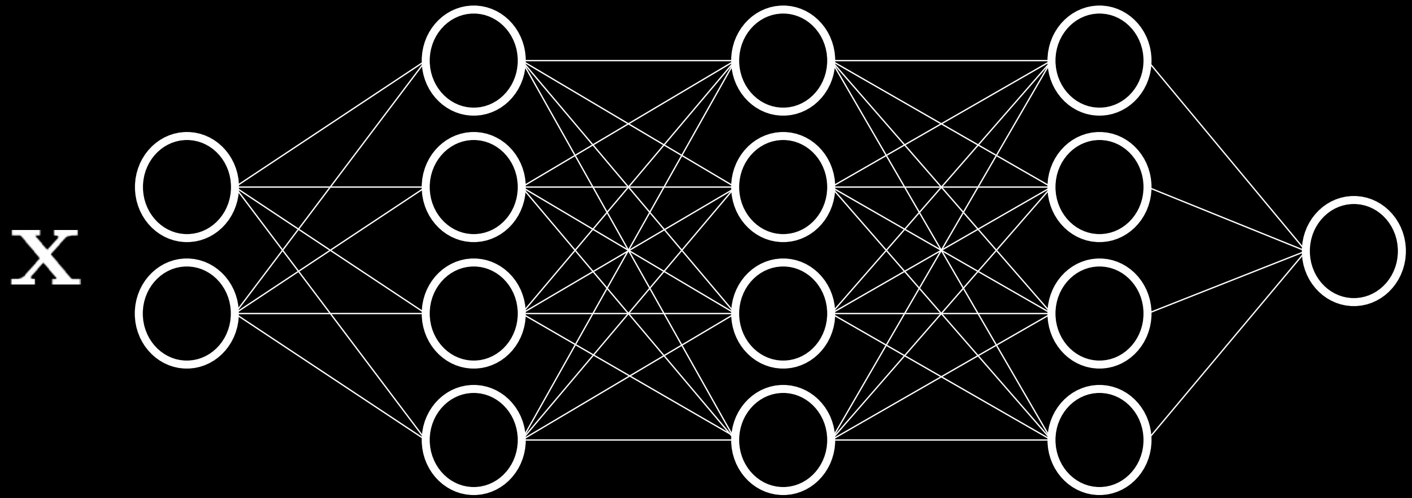


NeurIPS '20
CVPR '21
SIGGRAPH'21


Machine Learning/AI
& Optimization
Physics-Based
Visual Computing
Optics & Sensors
Imaging
&Computer Vision


Images
Shapes

Neural Networks as Signal Representations

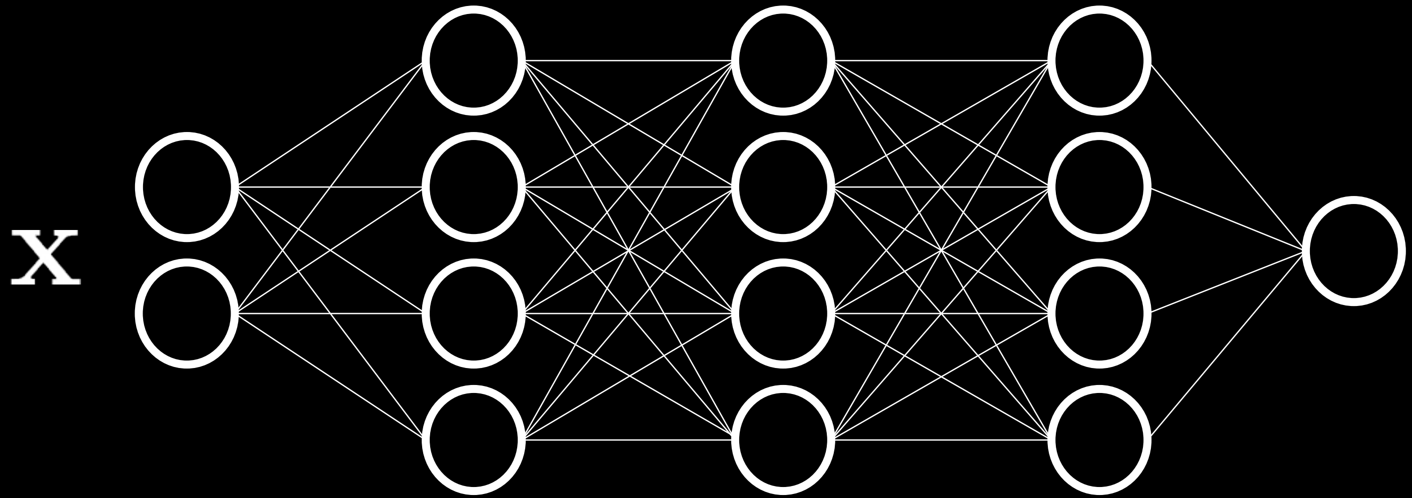


$\Phi(\mathbf{x}) : \mathbb{R}^2 \mapsto \mathbb{R}$

pixel coordinates

pixel values

Neural Networks as Signal Representations



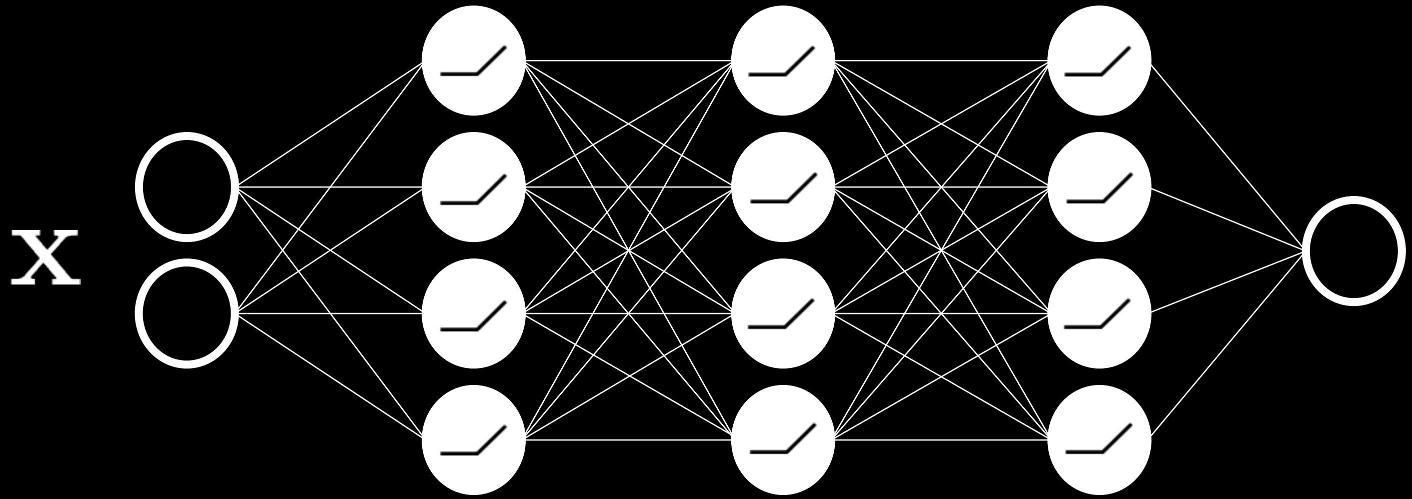
$$\Phi(\mathbf{x}) : \mathbb{R}^2 \mapsto \mathbb{R}$$

pixel coordinates

pixel values

- Agnostic to grid resolution
- Model memory scales with signal complexity
- Admits effective learning of priors
- Flexible, can be used with physics-based equations

Neural Networks as Signal Representations



$$\Phi(\mathbf{x}) : \mathbb{R}^2 \mapsto \mathbb{R}$$

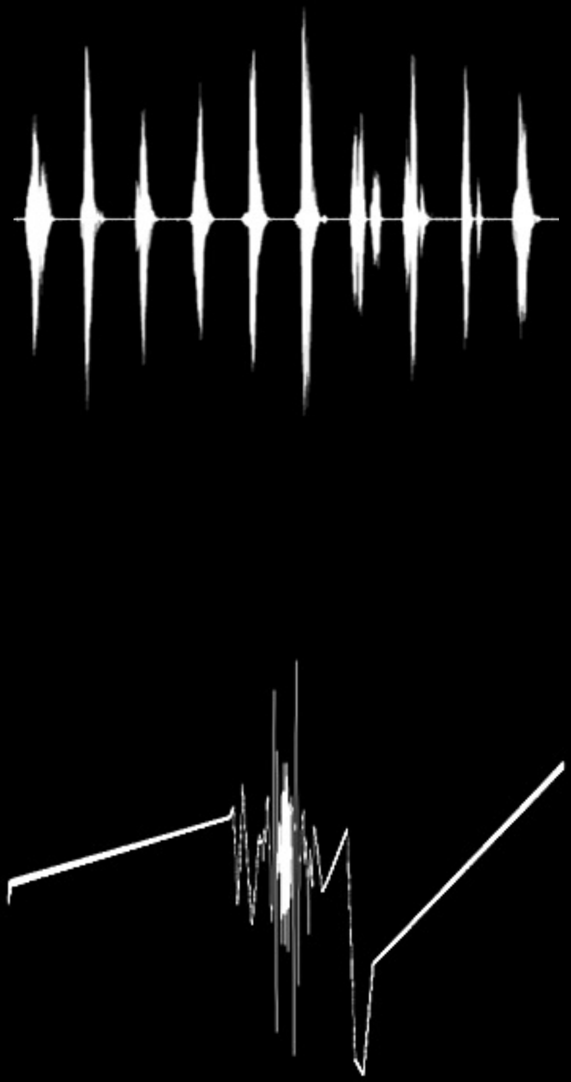
expected

actual

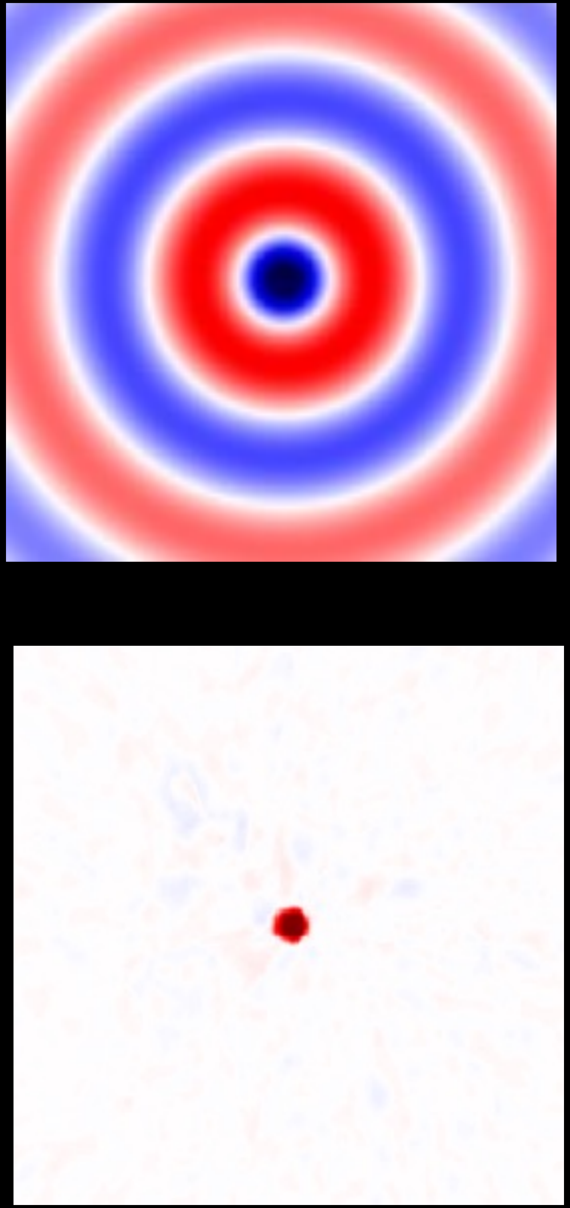
Images

Shapes

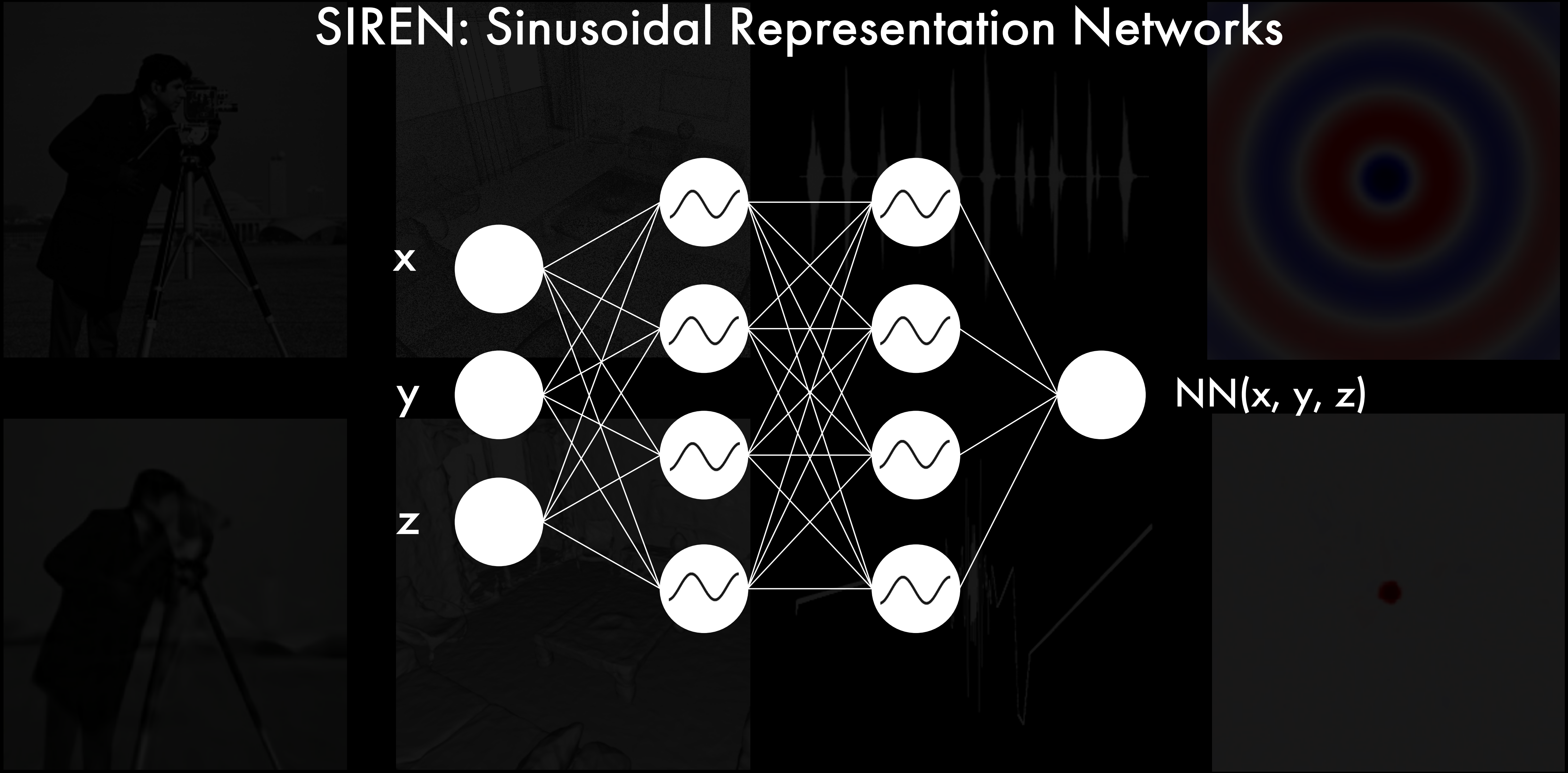
Audio



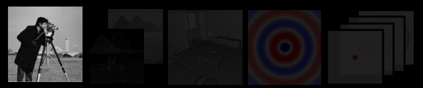
Quantities defined by a differential equation

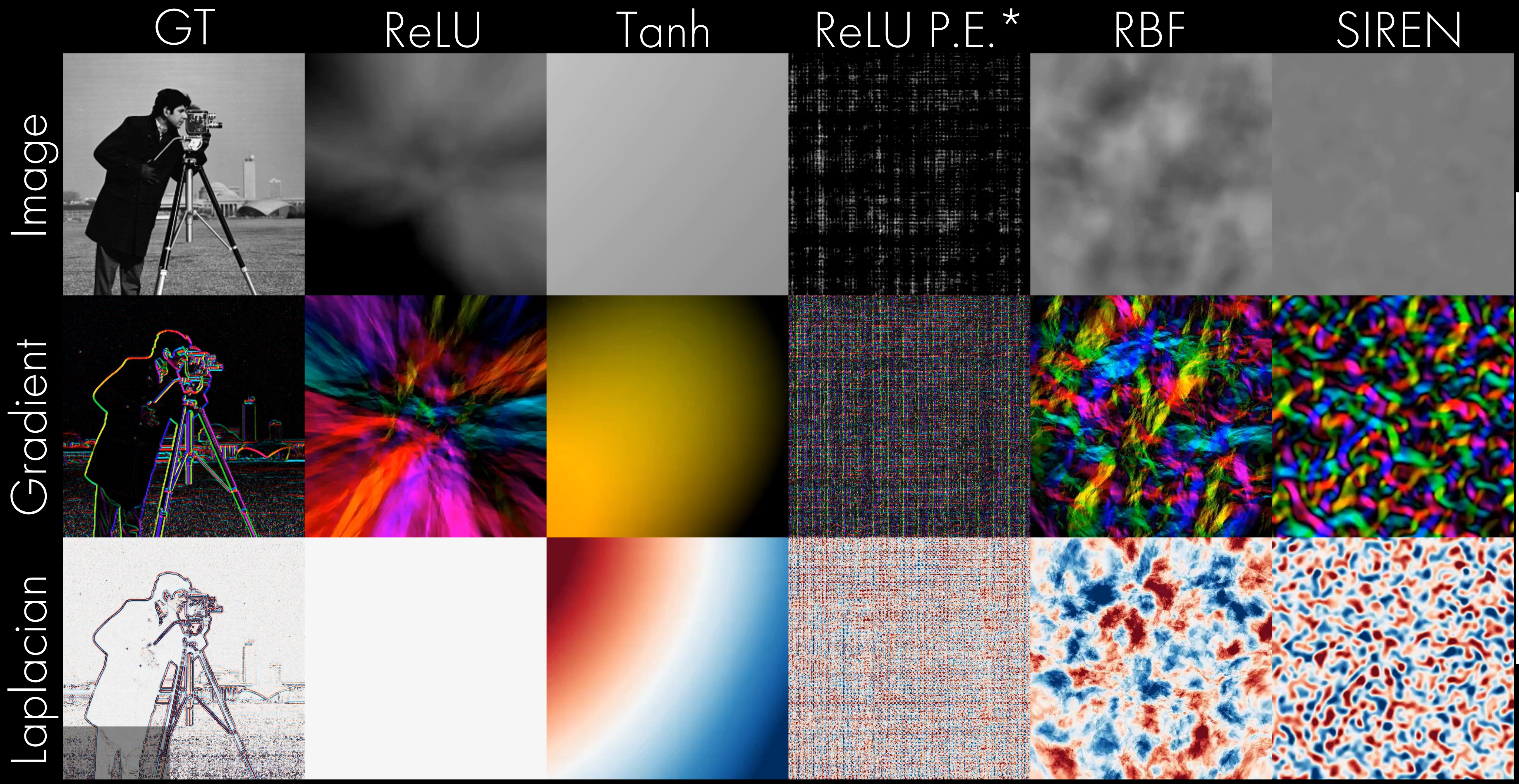


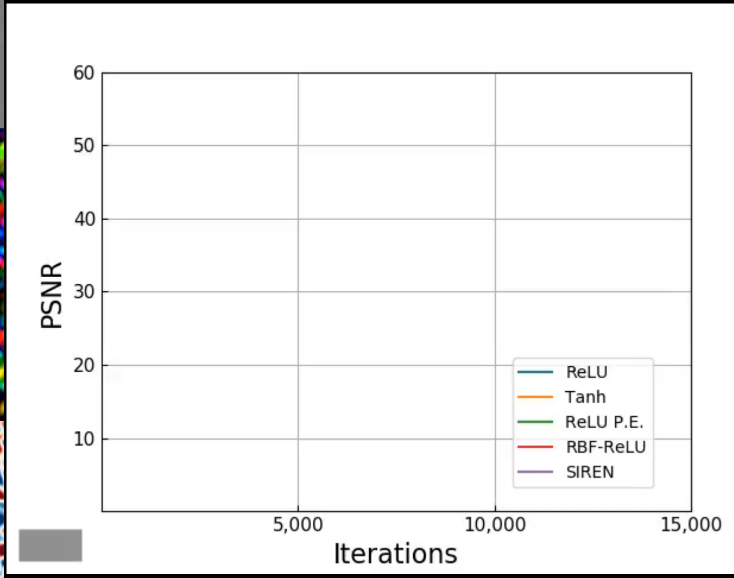
SIREN: Sinusoidal Representation Networks



Representing Images







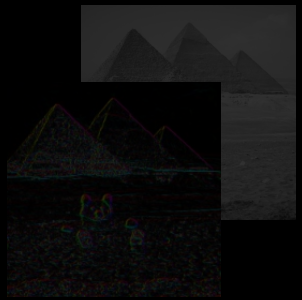
* Mildenhall et al. 2020

Images

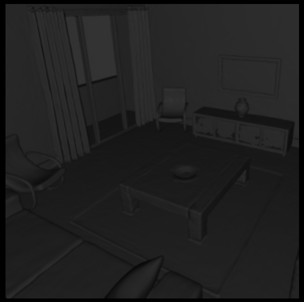
Audio

Videos

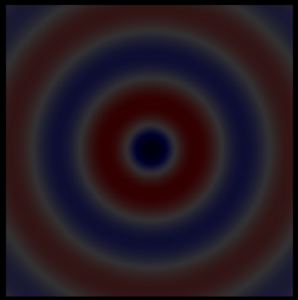
Poisson equation



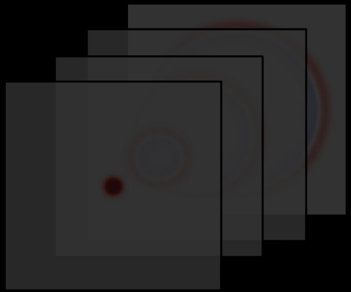
Eikonal equation



Helmholtz equation

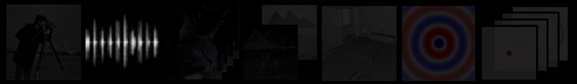


Wave equation

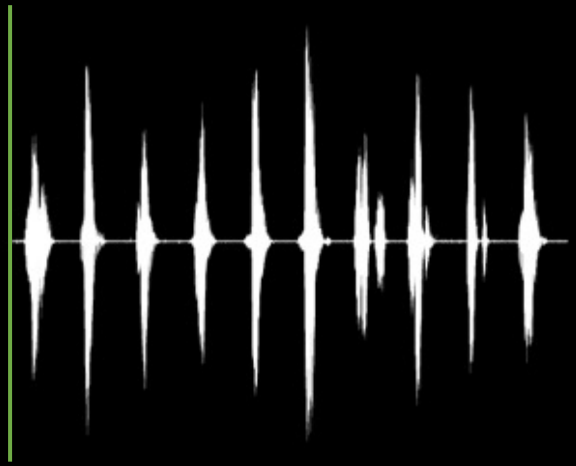


Input	Output supervised by	Implicit Formulation Find $\boldsymbol{\Phi}$ that minimizes $\boldsymbol{\mathcal{L}}$
$t \in \mathbb{R}$	$f(t) \in \mathbb{R}$	$\mathcal{L}_{\mathrm{audio}} = \int_{\Omega} \lVert \Phi(t) - f(t) \rVert \, \mathrm{d}t$

Representing Audio – Voice

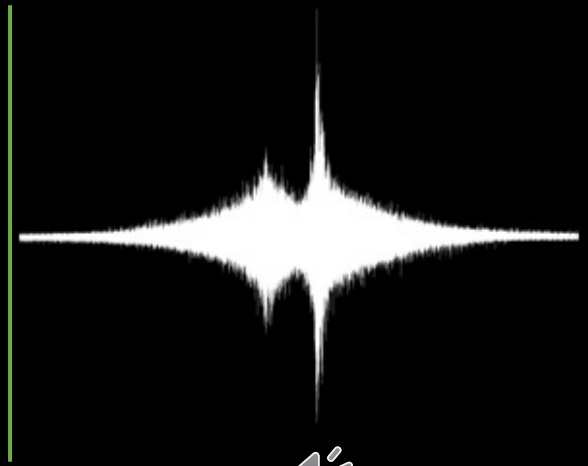


Ground Truth



ReLU MLP

ReLU w/ positional encoding

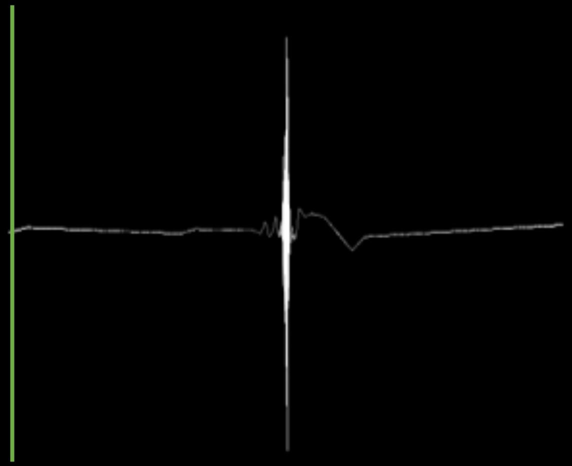


SIREN

Representing Audio – Music

Ground Truth

ReLU MLP



ReLU w/ positional encoding

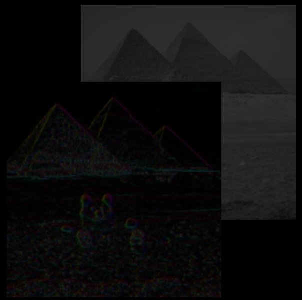
SIREN

Images

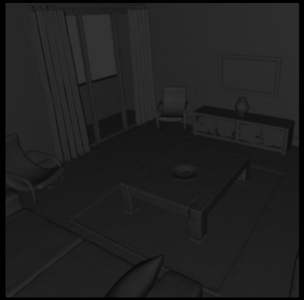
Audio

Videos

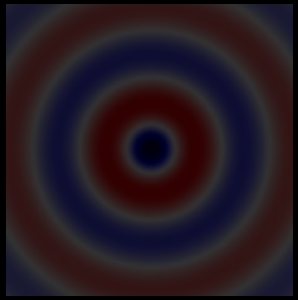
Poisson equation



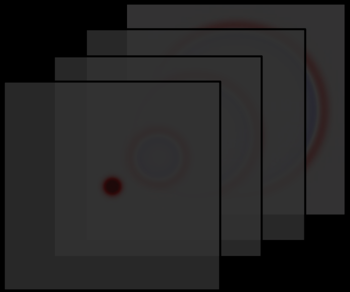
Eikonal equation



Helmholtz equation

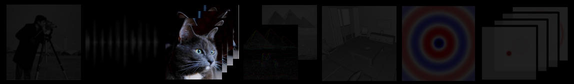


Wave equation



Input	Output supervised by	Implicit Formulation Find $\boldsymbol{\Phi}$ that minimizes $\boldsymbol{\mathcal{L}}$
$(\mathbf{x}, t) \in \mathbb{R}^3$	$f(\mathbf{x}, t) \in \mathbb{R}^3$	$\mathcal{L}_{\text{video}} = \int_{\Omega} \|\Phi(\mathbf{x}, t) - f(\mathbf{x}, t)\| \,\mathrm{d}\mathbf{x}\,\mathrm{d}t$
space-time coord.	*RGB value*	

Representing Video



Ground Truth

ReLU MLP

SIREN

Representing Video

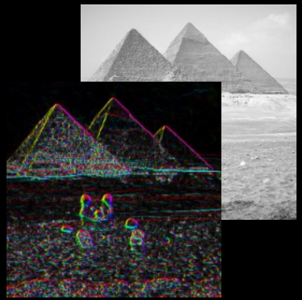
Ground Truth | ReLU MLP | SIREN

Images

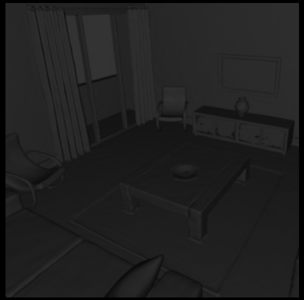
Audio

Videos

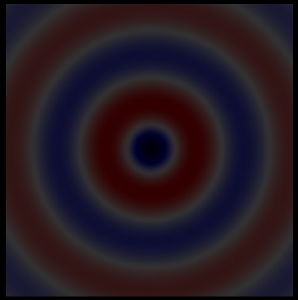
Poisson equation



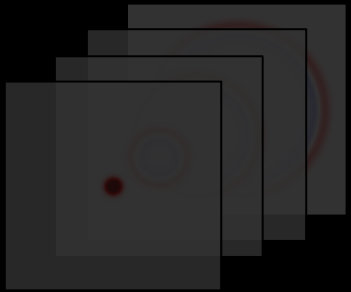
Eikonal equation



Helmholtz equation

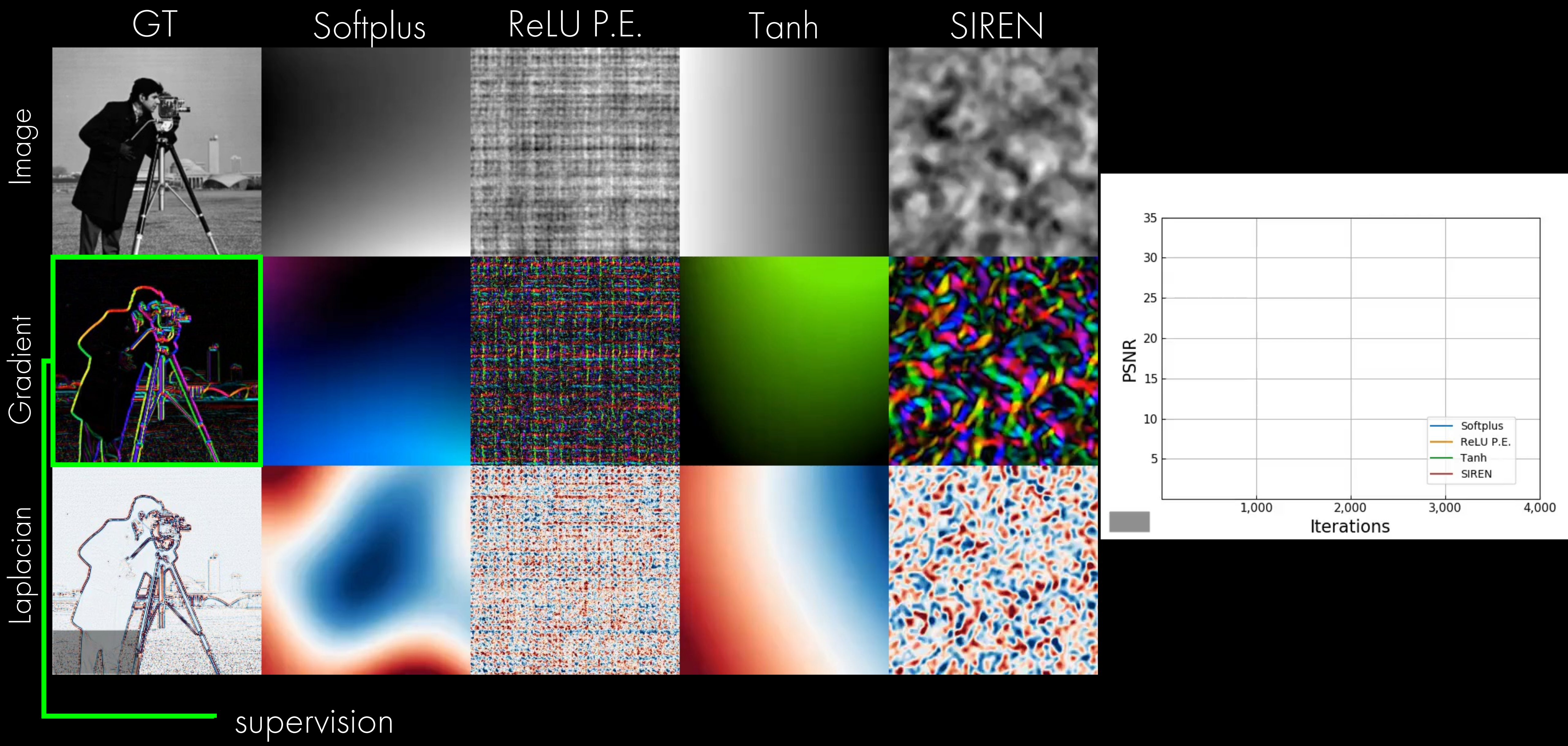


Wave equation

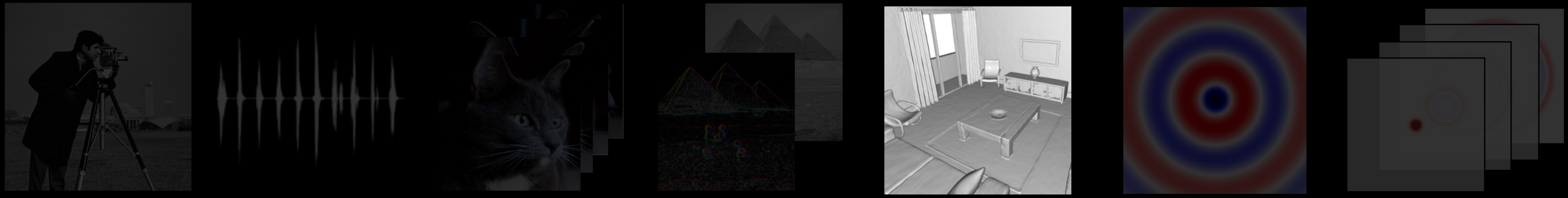


Input	Output supervised by	Implicit Formulation Find Φ that minimizes $\mathcal{L}$
$\mathbf{x} \in \mathbb{R}^2$	$f(\mathbf{x}) \in \mathbb{R}^2$	$\mathcal{L}_{\text{Poisson}} = \int_{\Omega} \lVert \nabla \Phi(\mathbf{x}) - \nabla f(\mathbf{x}) \rVert \, d\mathbf{x}$
spatial coord.	*gray level*	

Poisson's Equation

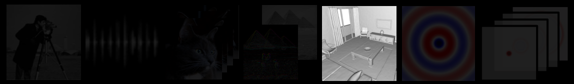


Images | Audio | Videos | Poisson equation | **Eikonal equation** | Helmholtz equation | Wave equation



Input	Output supervised by	Implicit Formulation Find Φ that minimizes $\mathcal{L}$
$\mathbf{x} \in \mathbb{R}^3$ *spatial coord.*	$f(\mathbf{x}) \in \mathbb{R}$ *signed distance*	$\mathcal{L}_{\text{Eikonal}} = \int_{\Omega_0} \lvert\Phi(\mathbf{x})\rvert + (1 - \langle \nabla\Phi(\mathbf{x}), \nabla f(\mathbf{x})\rangle)\, \mathrm{d}\mathbf{x} + \int_{\Omega} \lVert \lvert\nabla\Phi(\mathbf{x})\rvert - 1 \rVert \, \mathrm{d}\mathbf{x}$

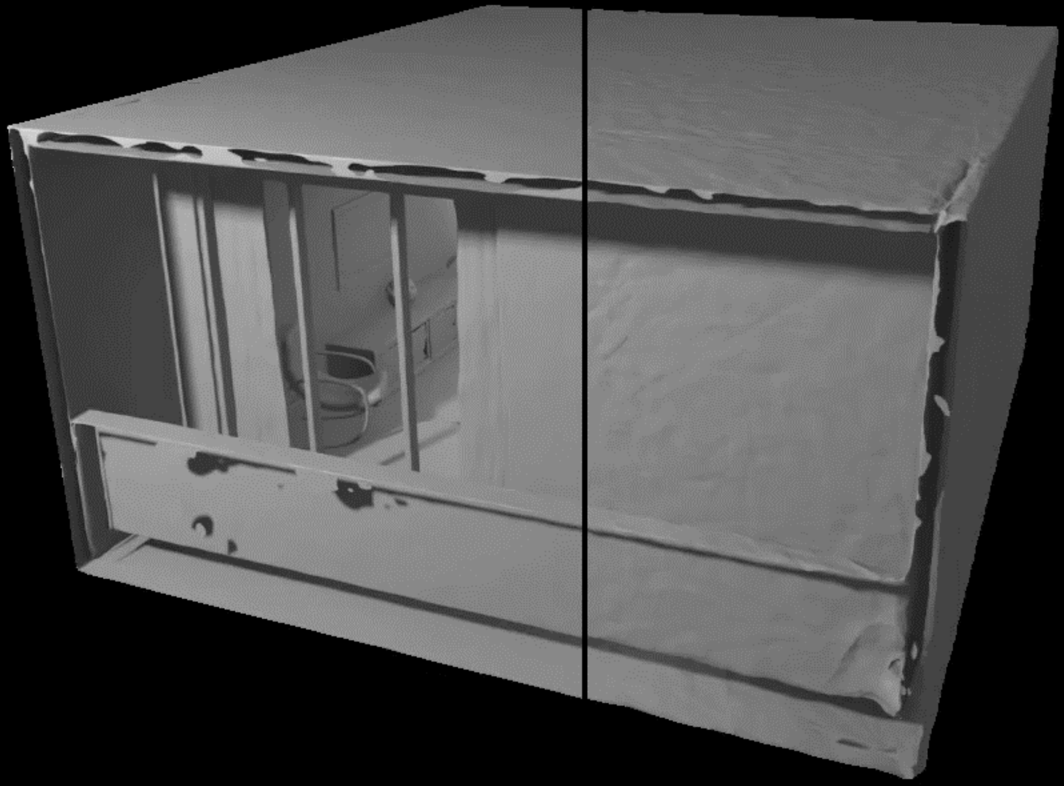
3D Shapes - solving the Eikonal equation



ReLU

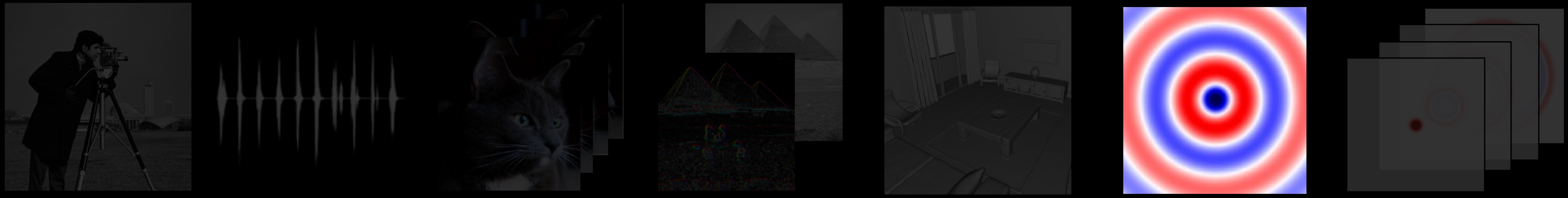
SIREN

5 layers,
256 hidden units



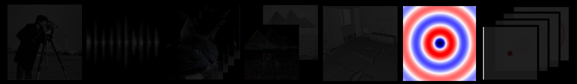
ReLU

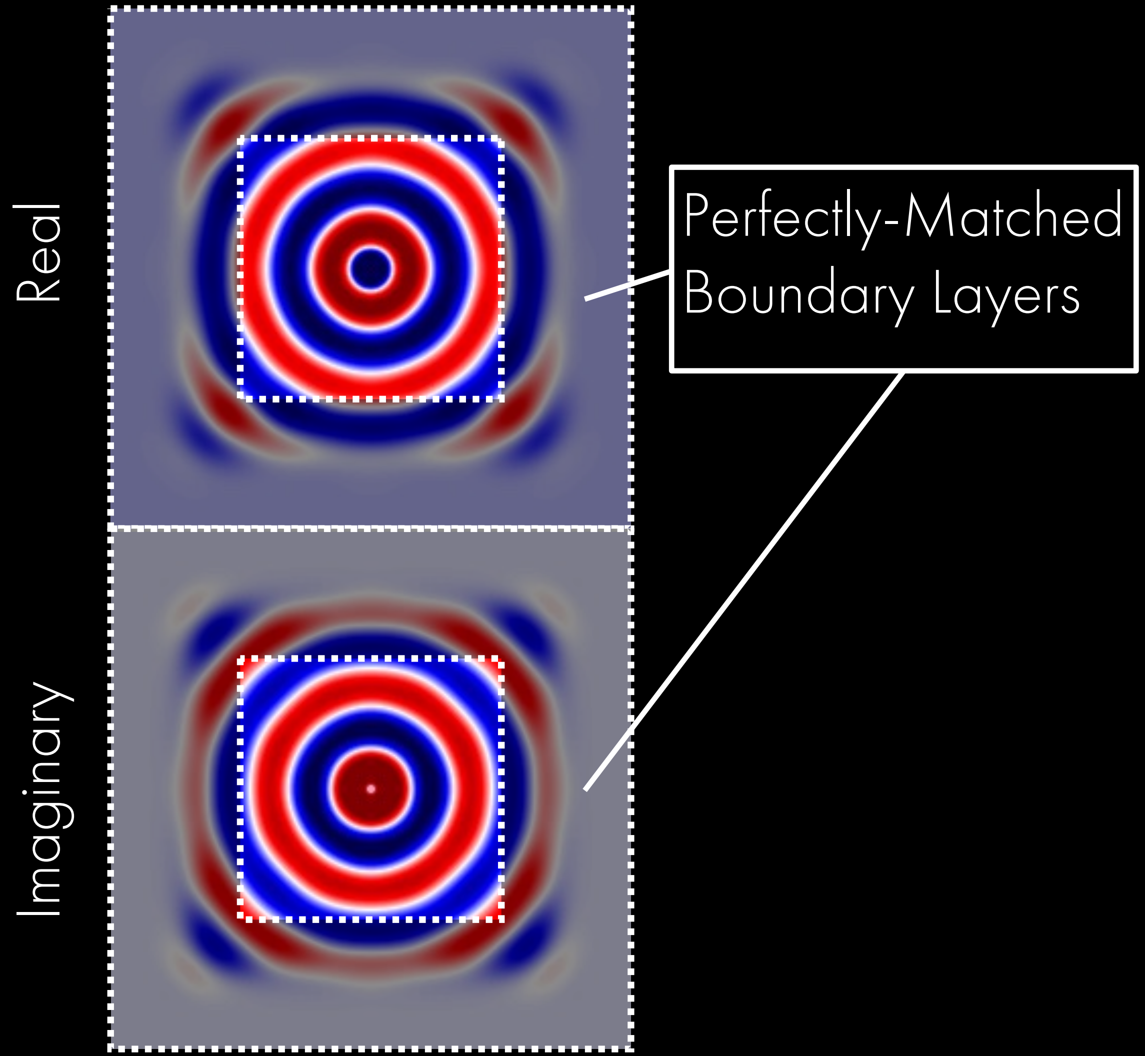
SIREN



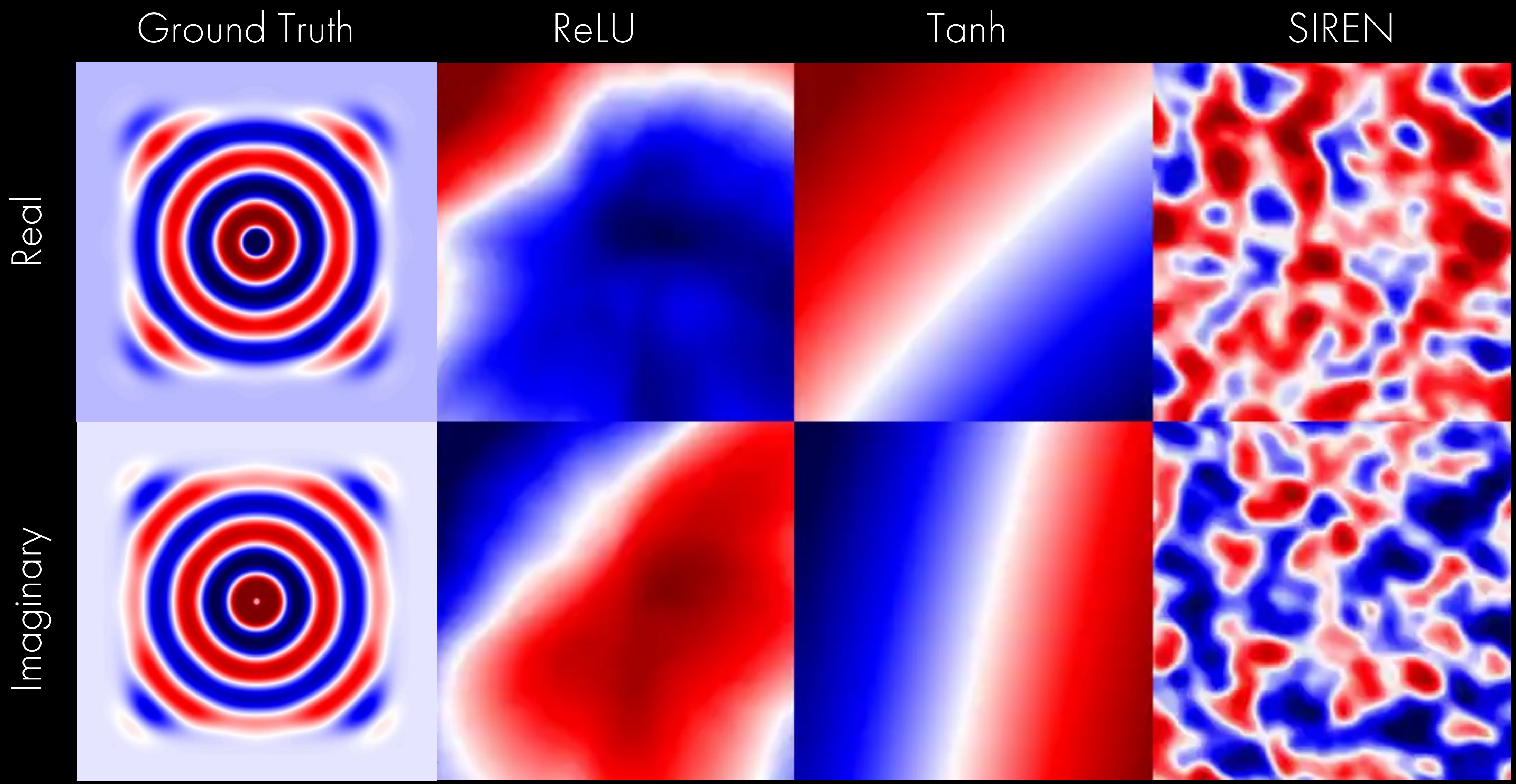
Input	Output supervised by	Implicit Formulation Find Φ that minimizes $\mathcal{L}$
$\mathbf{x} \in \mathbb{R}^2$	$f(\mathbf{x}) \in \mathbb{C}$	$\mathcal{L}_{\text{Helmholtz}} = \int_{\Omega} \|(\Delta + m(x)\omega^2)\Phi(\mathbf{x}) - f(\mathbf{x})\| \, d\mathbf{x}$
spatial coord.	*complex wave field*	

Solving the Helmholtz Equation

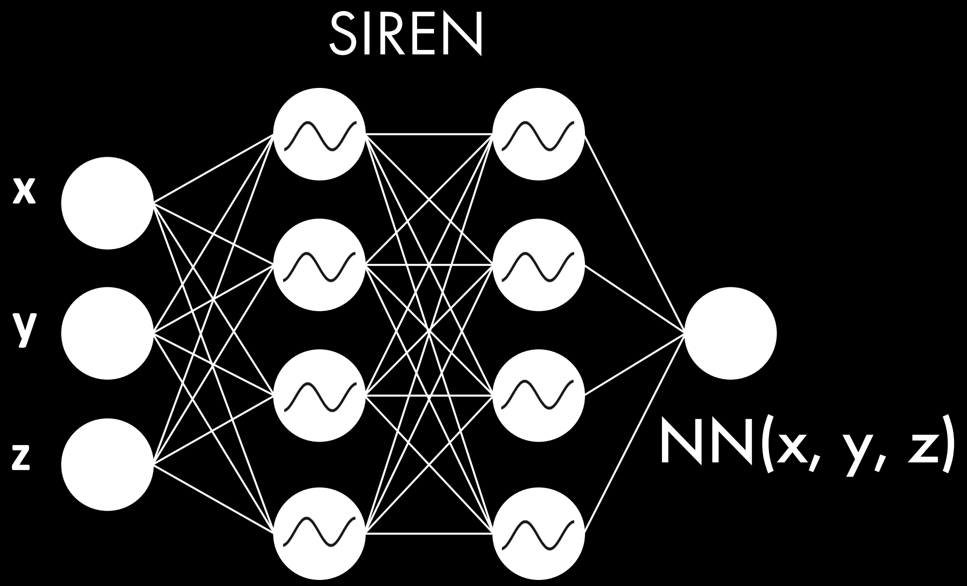


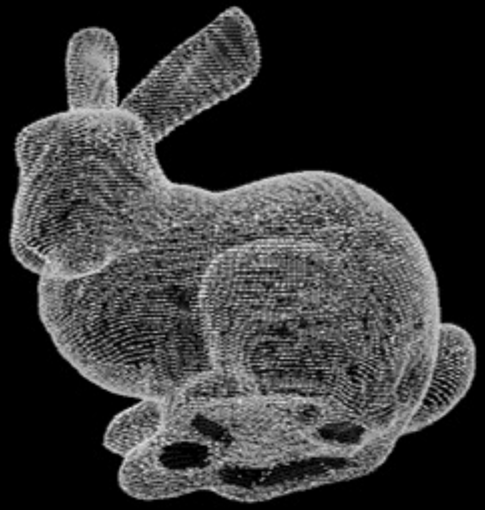


Solving the Helmholtz Equation

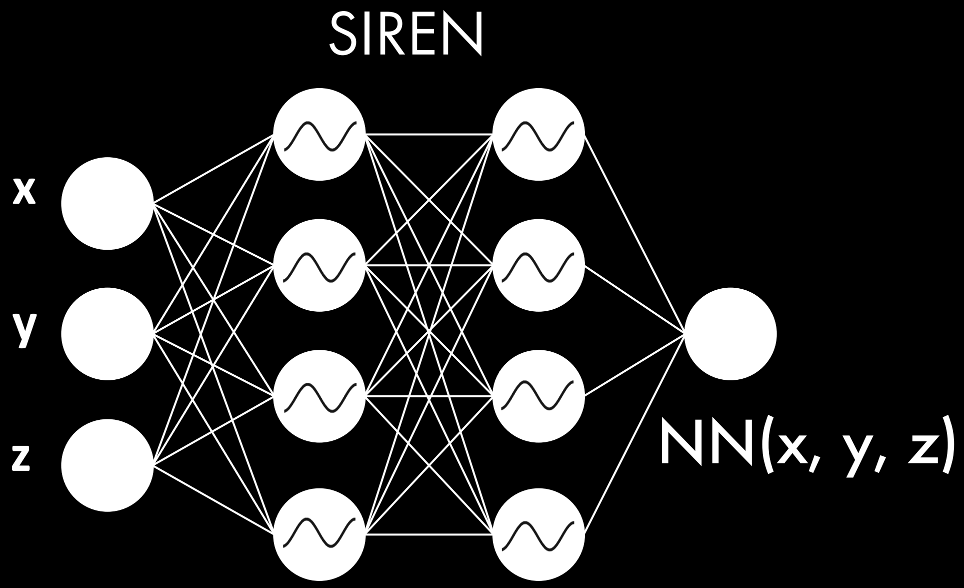


Like discrete grid or point clouds, SIREN is a data representation.





With a number of benefits.



Continuous, parametric (NN) function

Memory scale with signal complexity, independent of resolution

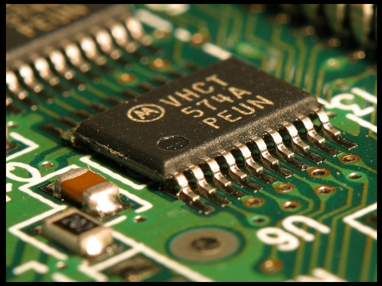
Can fit signals via first- and higher-order derivatives

Future Directions

Photon Interactions

time of flight
polarization
spectrum
coherence
angle
spatial statistics
...

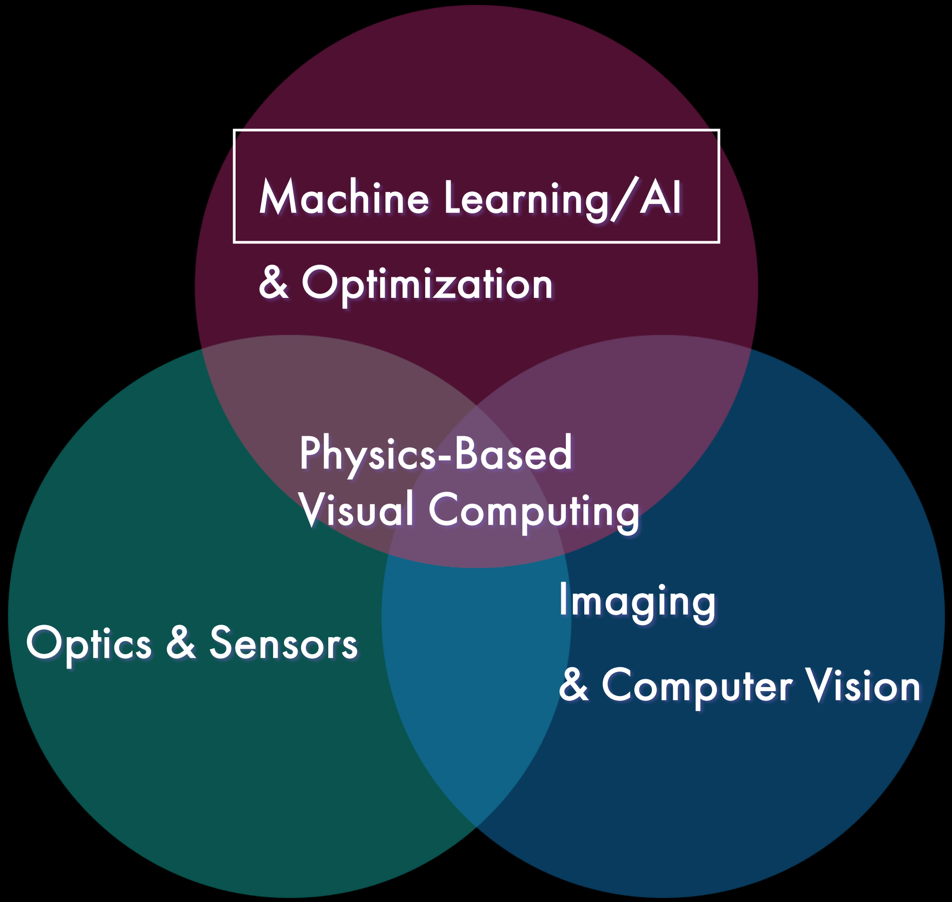
Optics, Sensors, Algorithms



"Superhuman" Visual Computing

3D geometry, lighting, reflectance, material properties, motion, segmentation, semantics, behavior, etc.

Future Directions



Future Directions

- How can we combine computational imaging with physics-based AI?
  - Need faster training times, more scalable architectures for large-scale signals

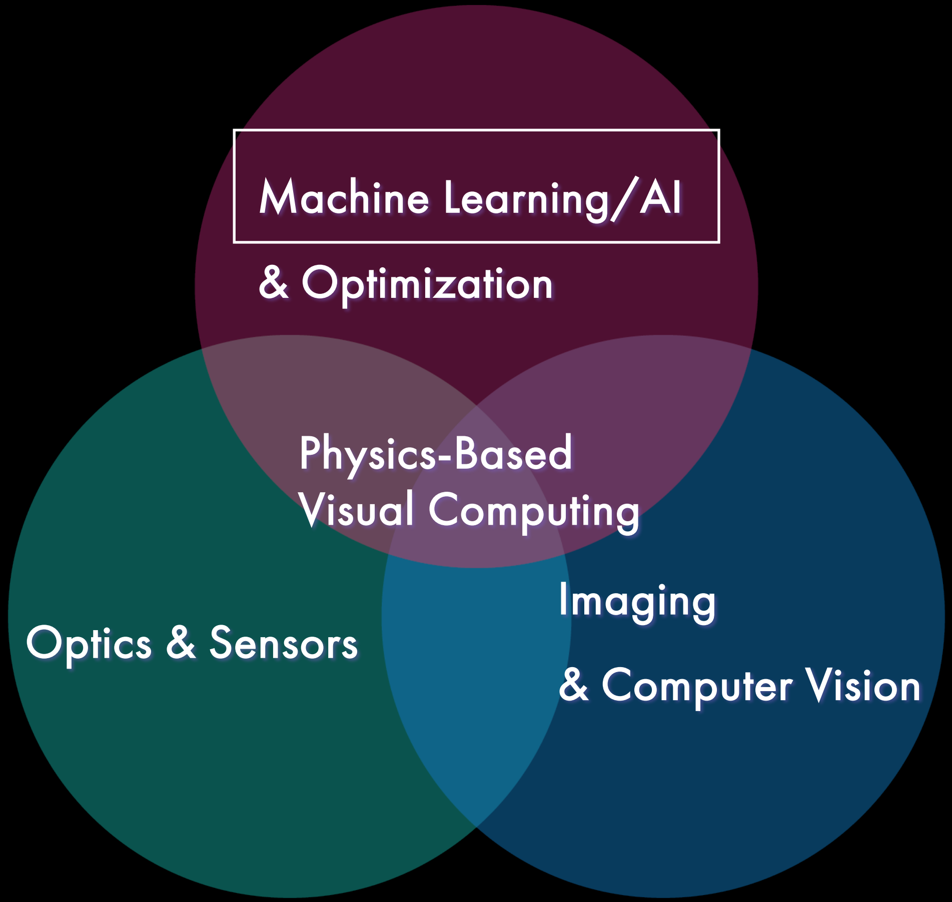
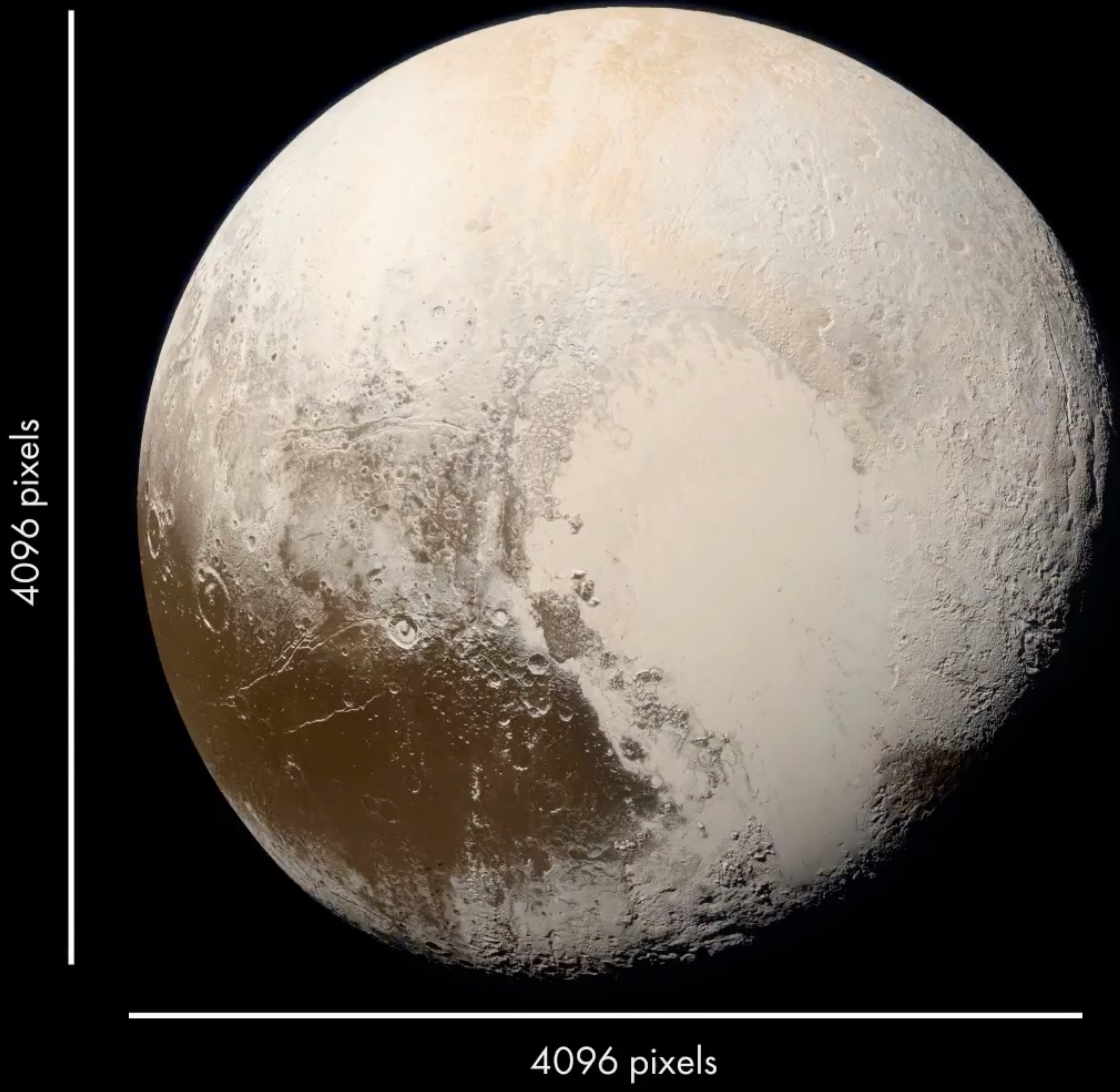


Image Fitting Example (16 MP)



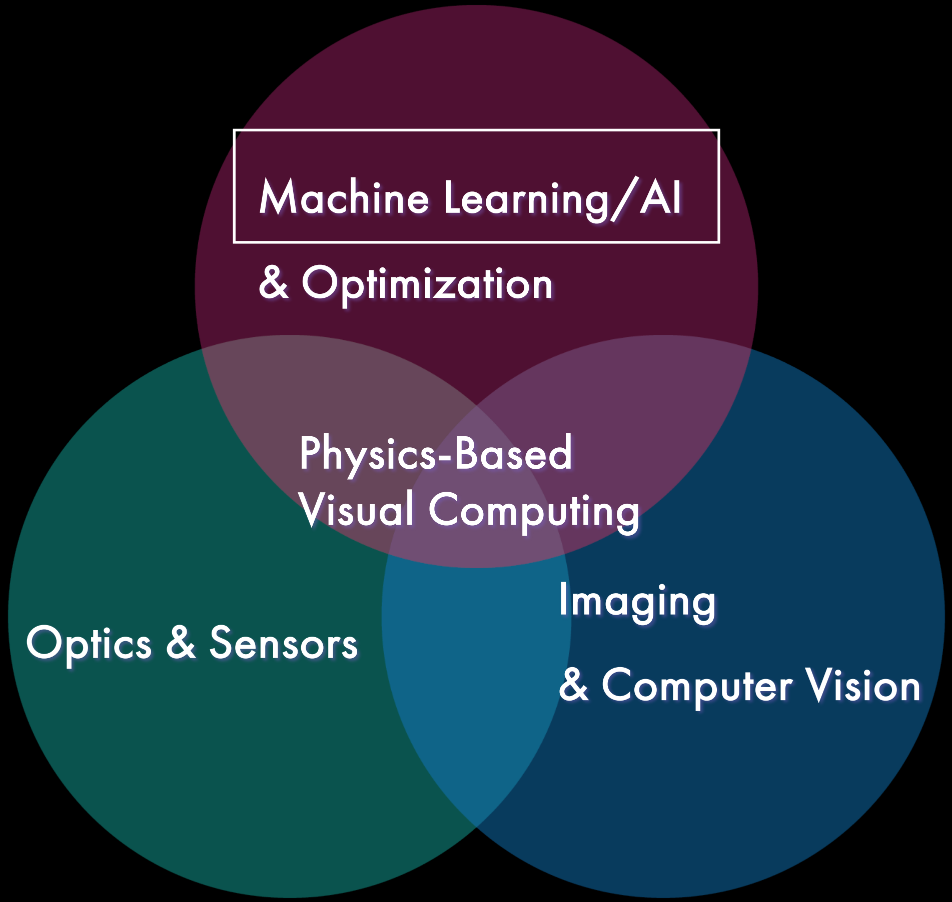
Pluto Image: NASA/JPL

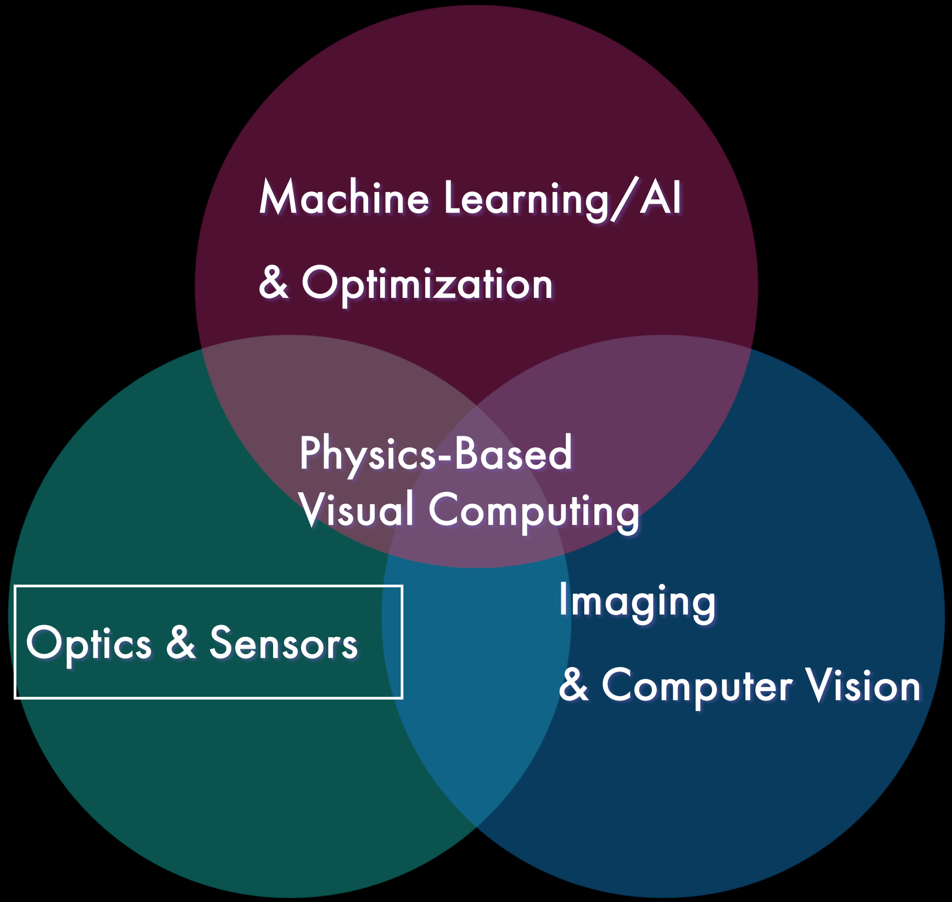
Gigapixel Image Fitting

ACORN
Output

Future Directions

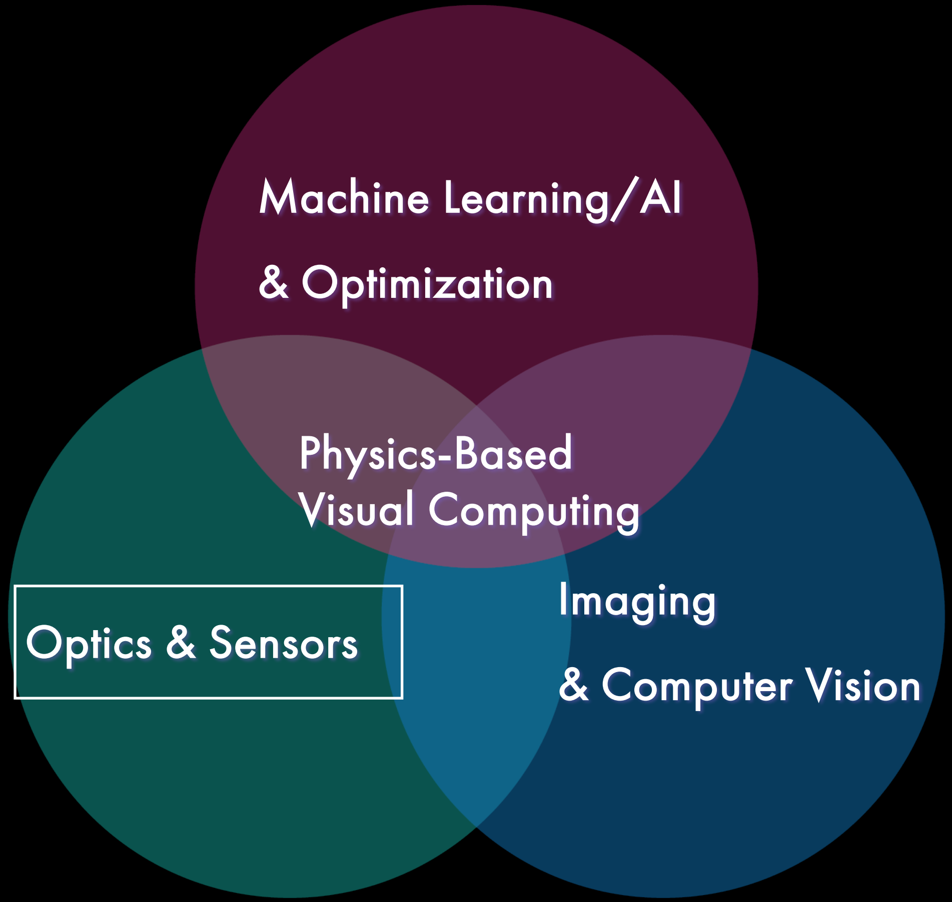
- How can we combine computational imaging with physics-based AI?
  - Need faster training times, more scalable architectures for large-scale signals
  - Generalization techniques to incorporate robust priors
  - Improve interpretability of representations
    (what about unsupervised input coordinates?)




Machine Learning/AI
& Optimization
Physics-Based
Visual Computing
Optics & Sensors
Imaging
& Computer Vision

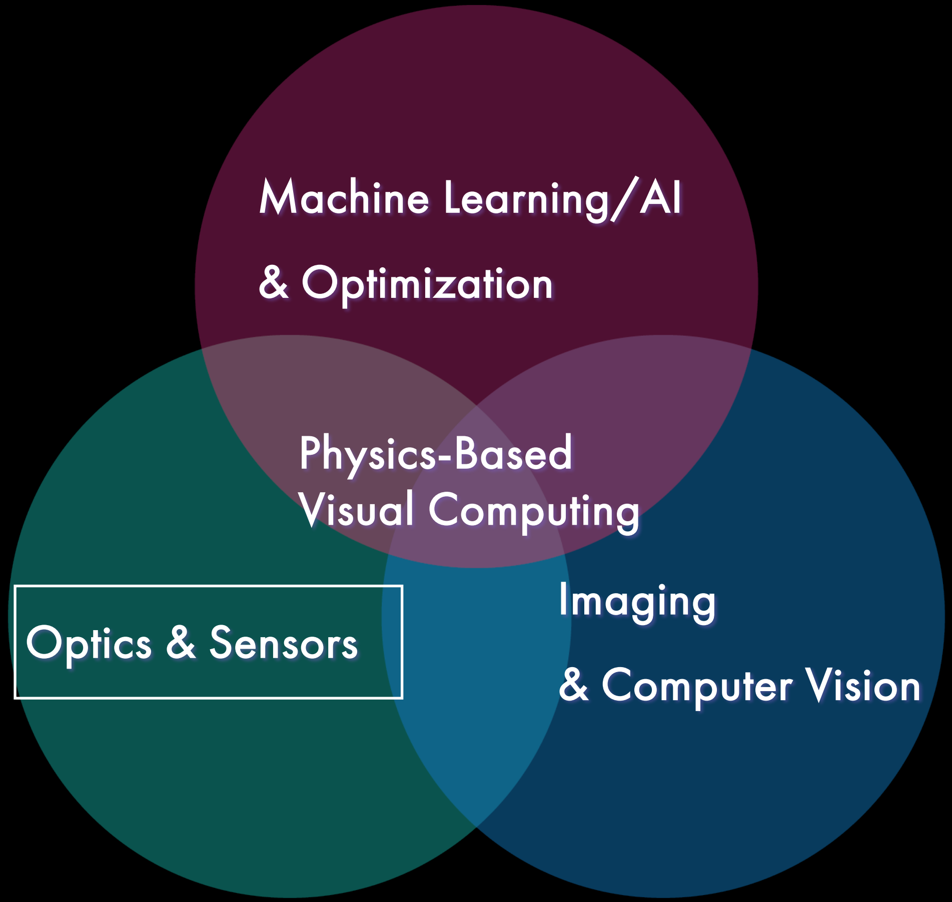
Future Directions

- Emerging "extreme" sensors
  - SPAD/jot arrays
    - ultra-low flux imaging
    - high-speed imaging
    - high-resolution imaging



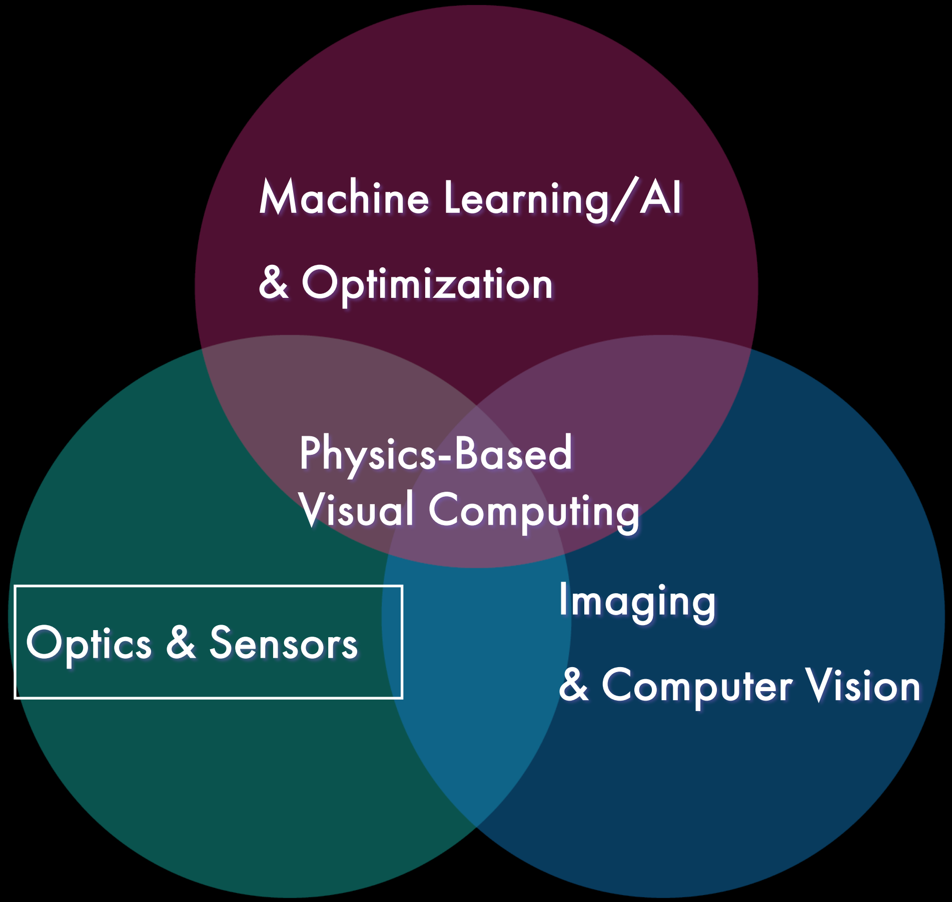
Future Directions

- Emerging "extreme" sensors
  - SPAD/jot arrays
    - ultra-low flux imaging
    - high-speed imaging
    - high-resolution imaging
  - Coherent LIDAR
    - micron-scale resolution
    - velocimetry
    - ambient rejection



Future Directions

- Emerging "extreme" sensors
  - SPAD/jot arrays
    - ultra-low flux imaging
    - high-speed imaging
    - high-resolution imaging
  - Coherent LIDAR
    - micron-scale resolution
    - velocimetry
    - ambient rejection
  - Sensor Fusion
    - LIDAR + radar + multiview stereo + acoustic

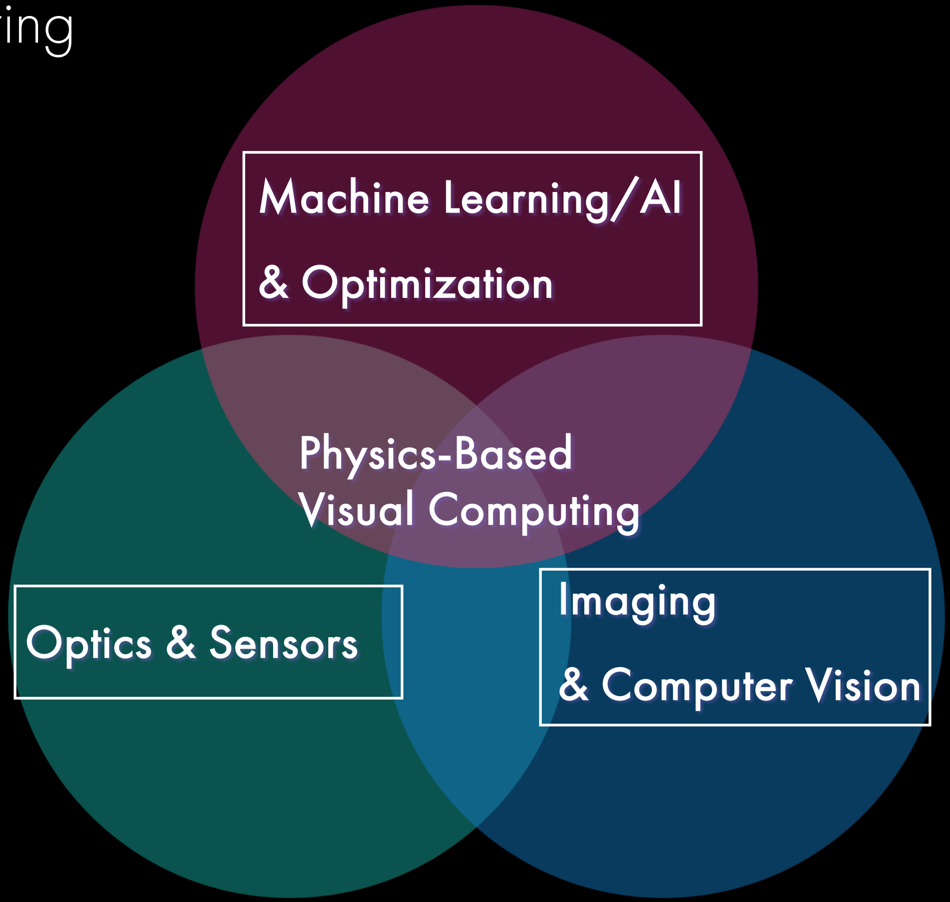


Future Directions

- Efficient solutions to radiative transfer
  - Biomedical imaging (micro-scale)
  - Robotics/remote sensing (macro-scale)
  - Many applications in computer vision, graphics, rendering

$$(\omega \cdot \nabla)L(\mathbf{x}, \omega) = -\sigma_t(\mathbf{x})L(\mathbf{x}, \omega) + L_e(\mathbf{x}, \omega)$$

$$+ \sigma_s(\mathbf{x}) \int_{\mathcal{S}^2} f_p(\mathbf{x}, \omega, \omega')L(\mathbf{x}, \omega')\mathrm{d}\omega'$$



Acknowledgments

Gordon Wetzstein

Alex Bergman

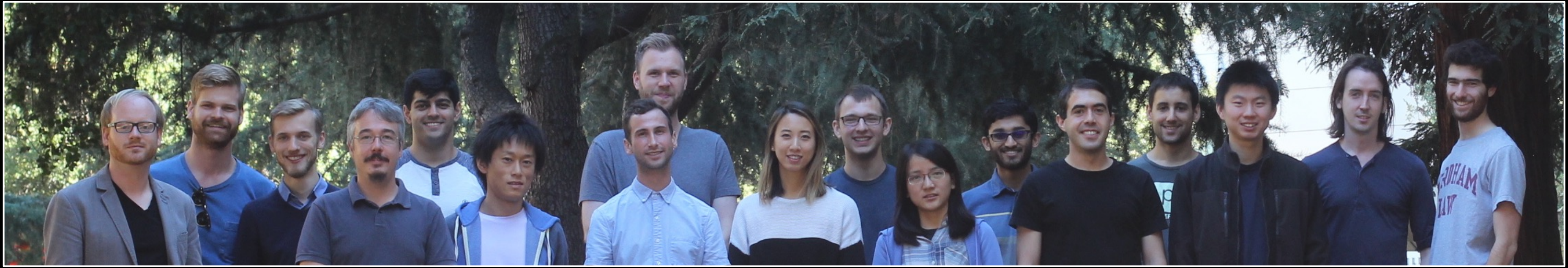
Julien Martel

Sean Young

Matt O’Toole

Chris Metzler

Vincent Sitzmann



SCI | STANFORD COMPUTATIONAL IMAGING LAB

intel

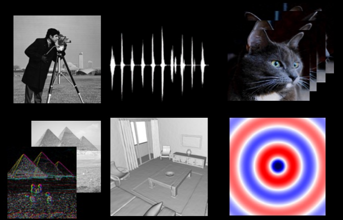
DARPA

ARL

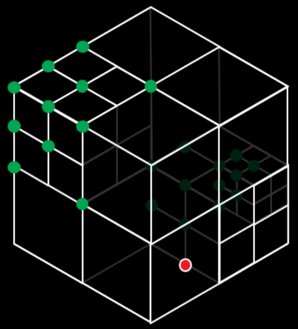
NSF

Computational Imaging at Toronto

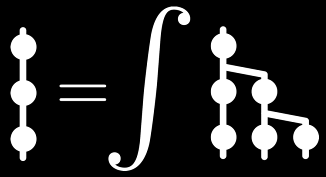
Machine Learning & 3D Vision



Physics-informed networks

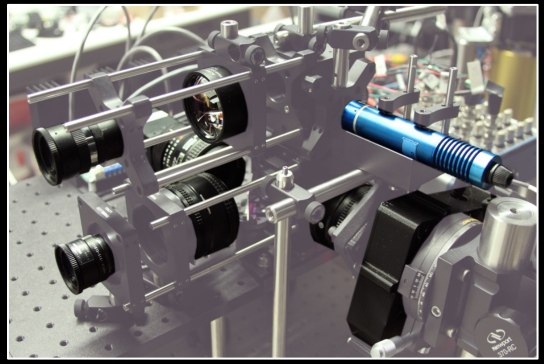


3D reconstruction

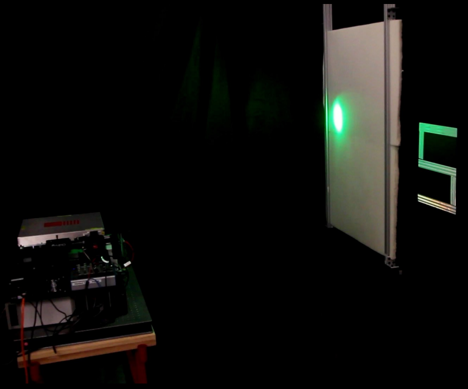


Neural rendering

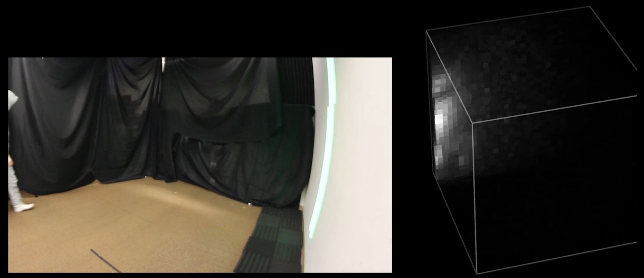
Computational Imaging



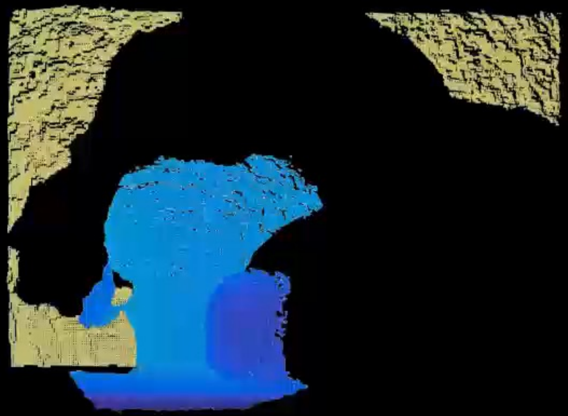
LIDAR



Imaging through scattering media



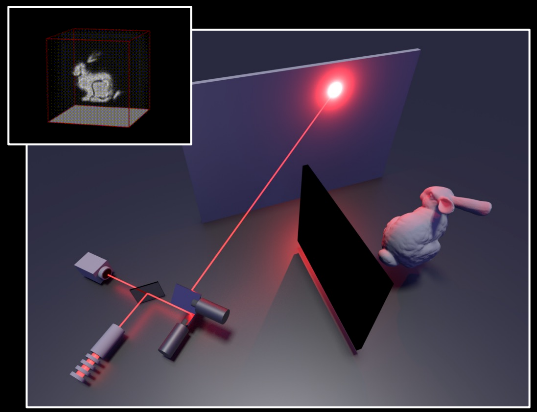
Non-line-of-sight imaging

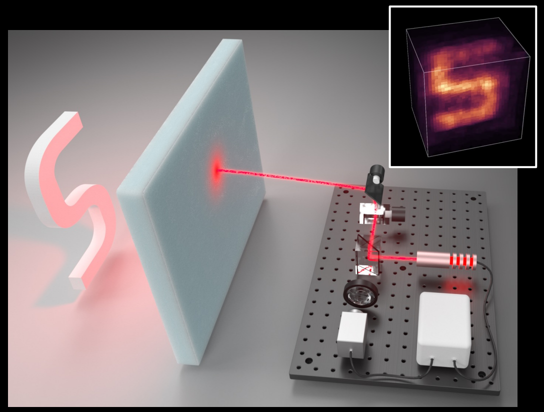


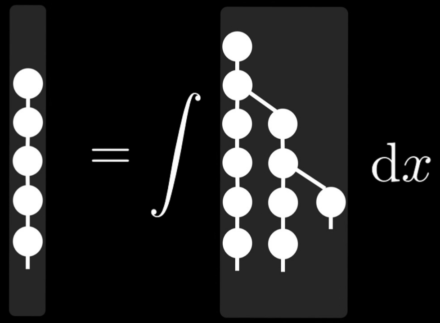
Single-photon imaging

Physics-Based Visual Computing for Efficient 3D Vision and Sensing

David B. Lindell
davidlindell.com
University of Toronto







References

Journal Publications

- **D. B. Lindell** and G. Wetzstein, "Three-dimensional imaging through scattering media based on confocal diffuse tomography," Nature Communications, vol. 11, no. 4517, 2020. [J10] C. A. Metzler, D. B.
- C. A. Metzler, **D. B. Lindell**, G. Wetzstein, "Keyhole imaging: Non-line-of-sight imaging and tracking of moving objects along a single optical path at long standoff distances," IEEE Trans. Comput. Imag., 2020, (Accepted).
- Z. Sun, **D. B. Lindell**, O. Solgaard, G. Wetzstein, "SPADnet: Deep RGB-SPAD sensor fusion assisted by monocular depth estimation," Optics Express, vol. 28, no. 10, pp. 14 948–14 962, 2020.
- F. Heide, M. O'Toole, K. Zang, **D. B. Lindell**, S. Diamond, G. Wetzstein, "Non-line-of-sight imaging with partial occluders and surface normals," ACM Trans. Graph., 2019
- **D. B. Lindell**, G. Wetzstein, M. O'Toole, "Wave-based non-line-of-sight imaging using fast f-k migration," ACM Trans. Graph. (SIGGRAPH), vol. 38, no. 4, 2019.
- F. Heide, S. Diamond, **D. B. Lindell**, G. Wetzstein, "Sub-picosecond photonefficient 3D imaging using single-photon sensors," Scientific Reports, vol. 8, no. 17726, 2018.
- **D. B. Lindell**, M. O'Toole, G. Wetzstein, "Single-photon 3D imaging with deep sensor fusion," ACM Trans. Graph. (SIGGRAPH), vol. 37, no. 4, 2018.
- M. O'Toole, **D. B. Lindell**, G. Wetzstein, "Confocal non-line-of-sight imaging based on the light cone transform," Nature, vol. 555, no. 7696, p. 338, 2018.

Conference Publications

- W. Bergman, **D. B. Lindell**, G. Wetzstein, "Deep adaptive LiDAR: End-toend optimization of sampling and depth completion at low sampling rates," in IEEE International Conference on Computational Photography (ICCP), 2020.
- M. Nishimura, **D. B. Lindell**, C. Metzler, G. Wetzstein, "Disambiguating monocular depth estimation with a single transient," in European Conference on Computer Vision (ECCV), 2020.
- V. Sitzmann, J. N. Martel, A. W. Bergman, **D. B. Lindell**, G. Wetzstein, "Implicit neural representations with periodic activation functions," in Advances in Neural Information Processing Systems (NeurIPS), 2020, (Oral).
- S. I. Young, **D. B. Lindell**, B. Girod, D. Taubman, G. Wetzstein, "Non-line-ofsight surface reconstruction using the directional light-cone transform," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, (Oral).
- **D. B. Lindell**, G. Wetzstein, V. Koltun, "Acoustic non-line-of-sight imaging," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, (Oral).
- **D. B. Lindell**, M. O'Toole, G. Wetzstein, "Towards transient imaging at interactive rates with single-photon detectors," in IEEE International Conference on Computational Photography (ICCP), 2018.
- M. O'Toole, F. Heide, **D. B. Lindell**, K. Zang, S. Diamond, G. Wetzstein, "Reconstructing transient images from single-photon sensors," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, (Spotlight).