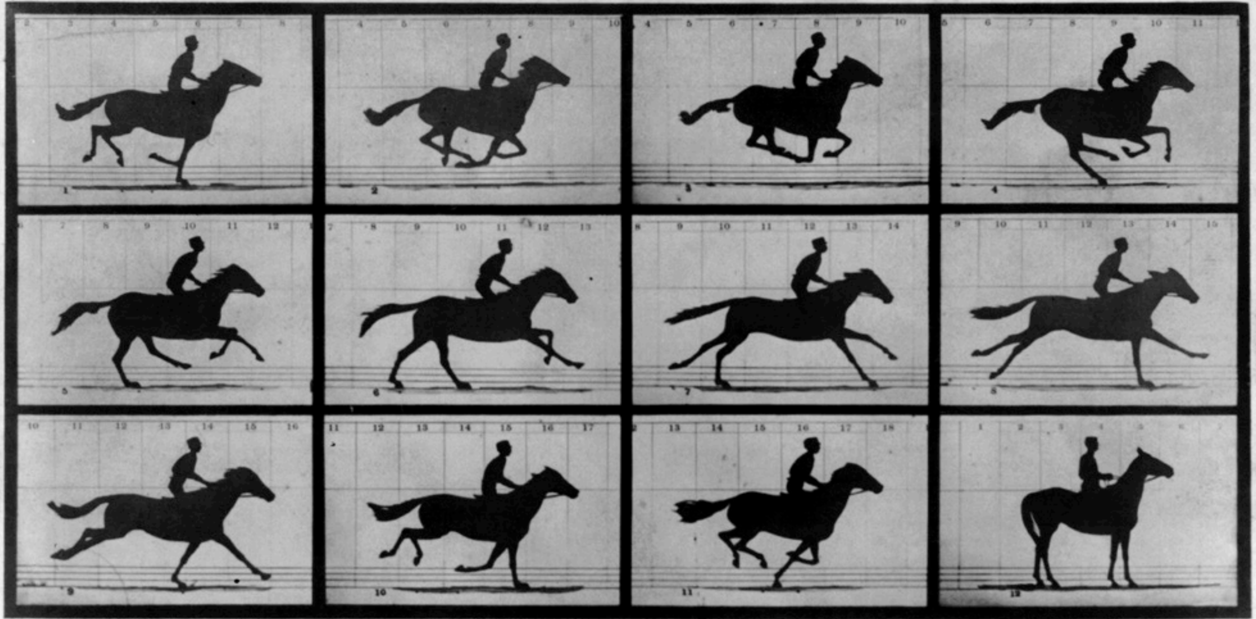
Computational Single-photon Imaging

Gordon Wetzstein
Stanford University
ISSW 2/28/2018

www.computationalimaging.org



MORSE'S Gallery, 417 Montgomery St., San Francisco.

THE HORSE IN MOTION.

Illustrated by

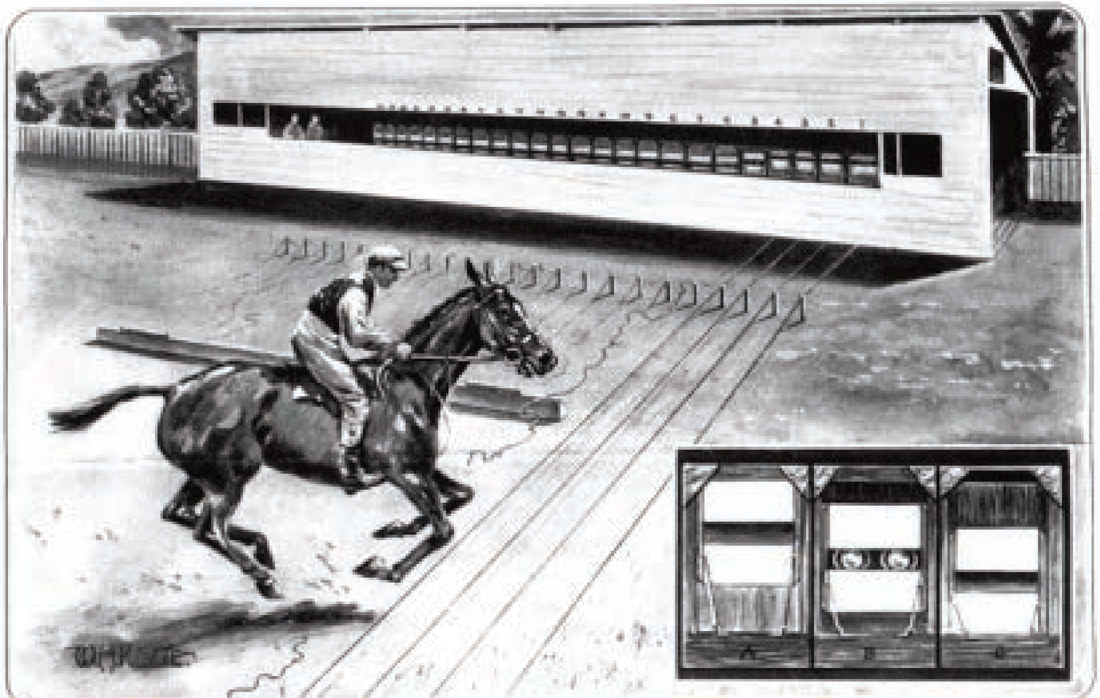
MUYBRIDGE.

AUTOMATIC ELECTRO-PHOTOGRAPH.

"SALLIE GARDNER," owned by LELAND STANFORD; running at a 1.40 gait over the Palo Alto track, 19th June, 1878.

The negatives of these photographs were made at intervals of twenty-seven inches of distance, and about the twenty-fifth part of a second of time; they illustrate consecutive positions assumed in each twenty-seven inches of progress during a single stride of the mare. The vertical lines were twenty-seven inches apart; the horizontal lines represent elevations of four inches each. The exposure of each negative was less than the two-thousandth part of a second.

Muybridge’s Multi-Camera Array at Stanford



Computational Imaging

Computational
Cameras

optics

+

sensing

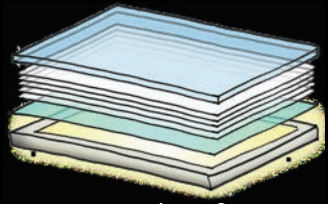
+

computation

Computational
Displays

computation

+



optics &
electronics

+

human visual
system

Computational Imaging

Computational Cameras

optics

+

sensing

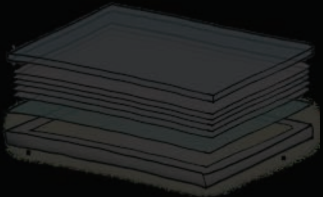
+

computation

Computational Displays

computation

+



optics & electronics

+

human visual system

Computational Imaging

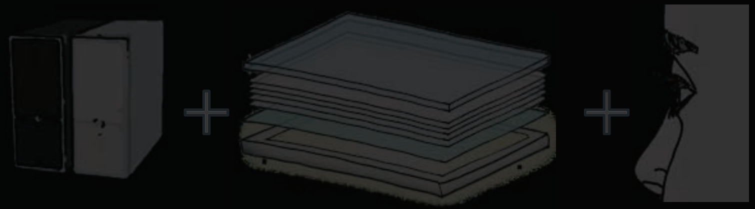
Computational Cameras

HDR Imaging [Debevec, Nayar, …]

Super-resolution [Baker, …]

Light Fields [Levoy, …]

Computational Displays



computation + optics & electronics + human visual system

Computational Imaging

Computational Cameras

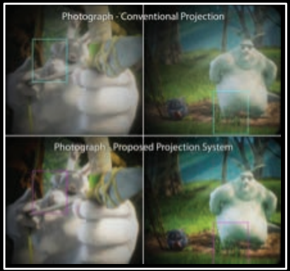
HDR Imaging [Debevec, Nayar, …]

Super-resolution [Baker, …]

Light Fields [Levoy, …]

Computational Displays

HDR Display [Seetzen, …]



Super-resolution [Hirsch, Heide, …]

Light Fields [Wetzstein, …]

Computational Light Transport

Computational
Cameras

+

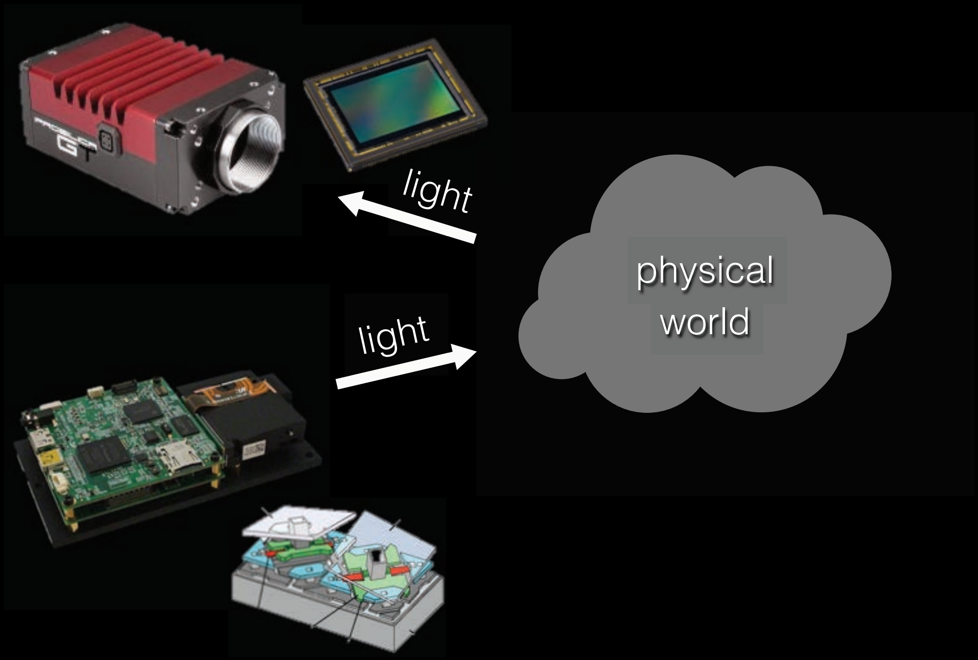
Computational
Displays

Computational Light Transport

Computational
Cameras

+

Computational
Displays

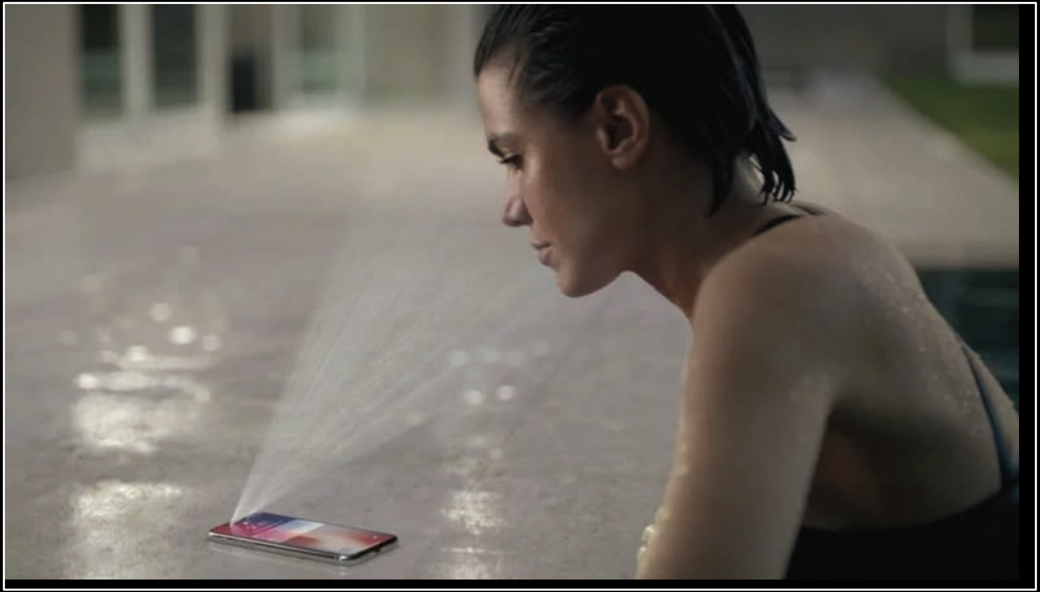


3D Imaging for Autonomous Vehicles

LIDAR (light detection and ranging)
Velodyne VLS-128

3D Imaging for Smartphones

iPhone X

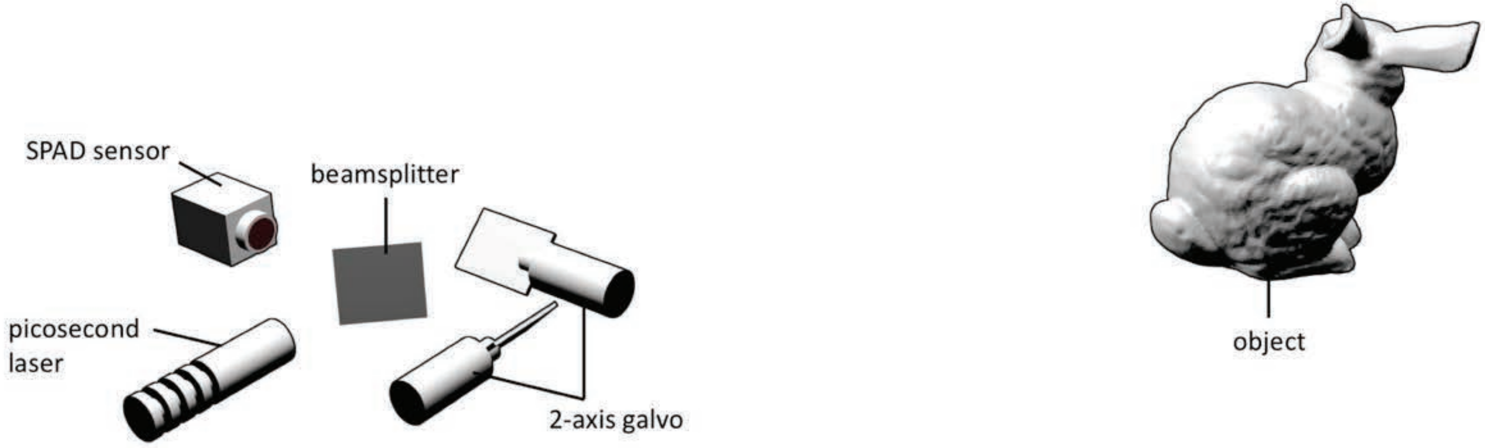


3D Imaging for VR/AR

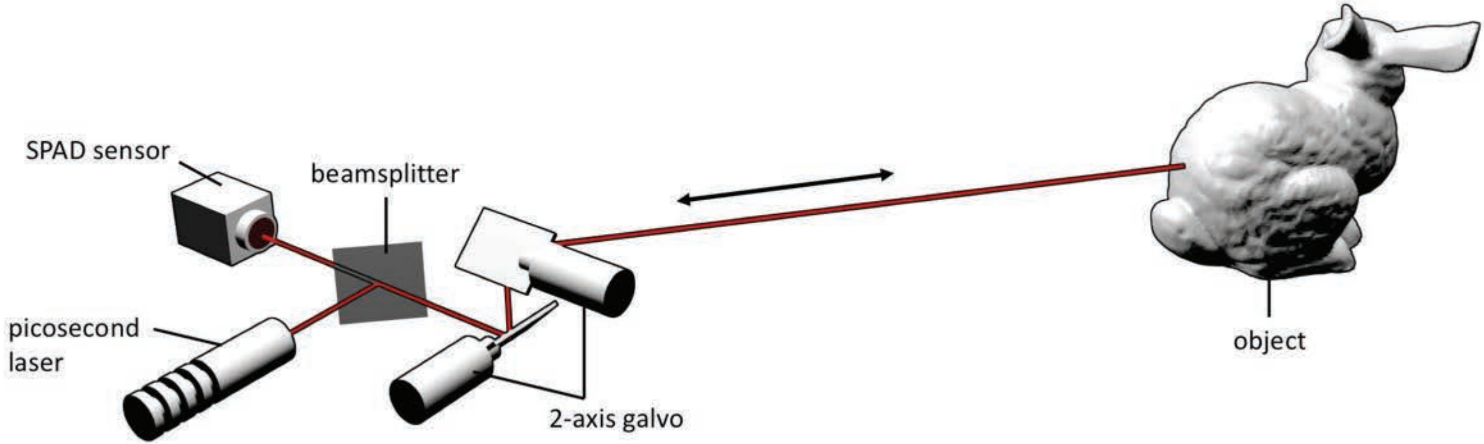
HTC Vive

HTC Vive Lighthouse

Direct Time-of-Flight 3D Imaging



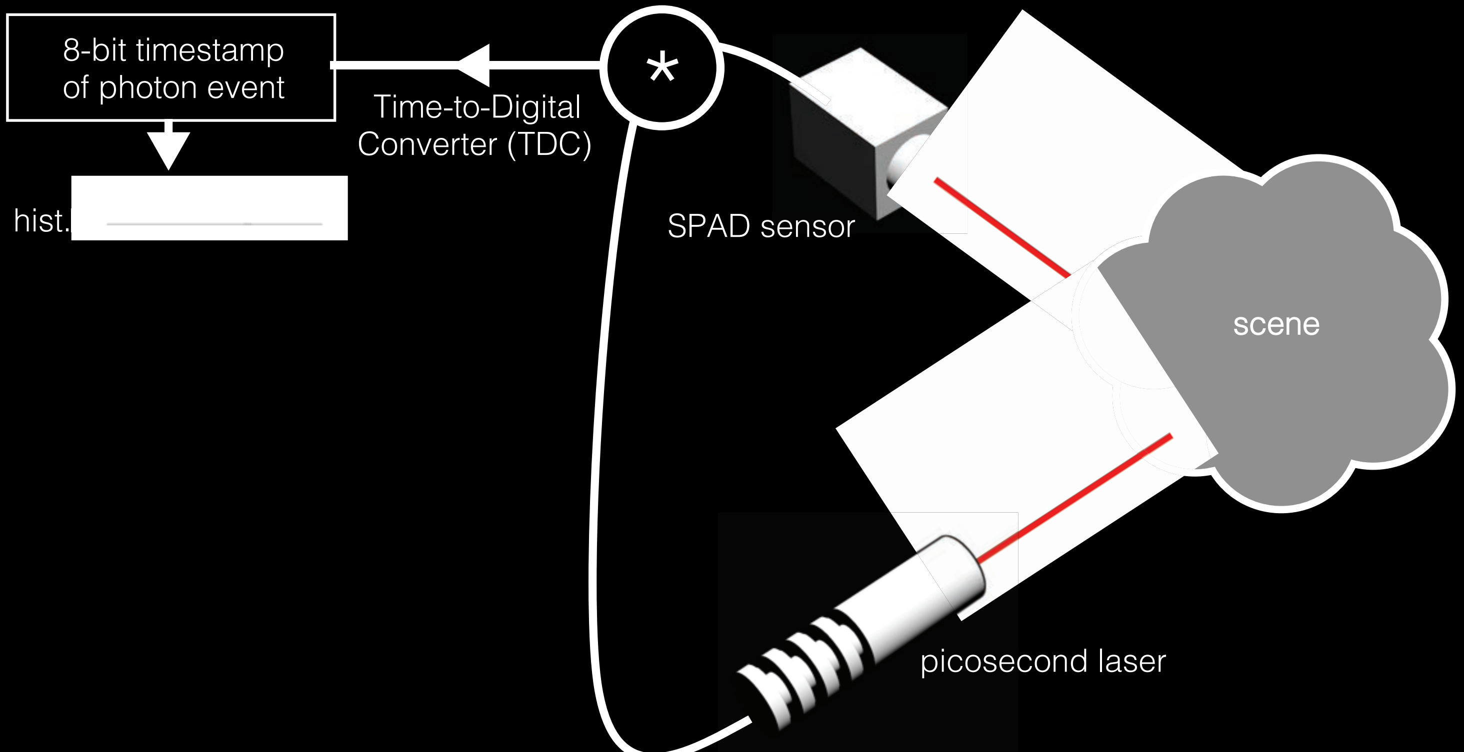
Direct Time-of-Flight 3D Imaging



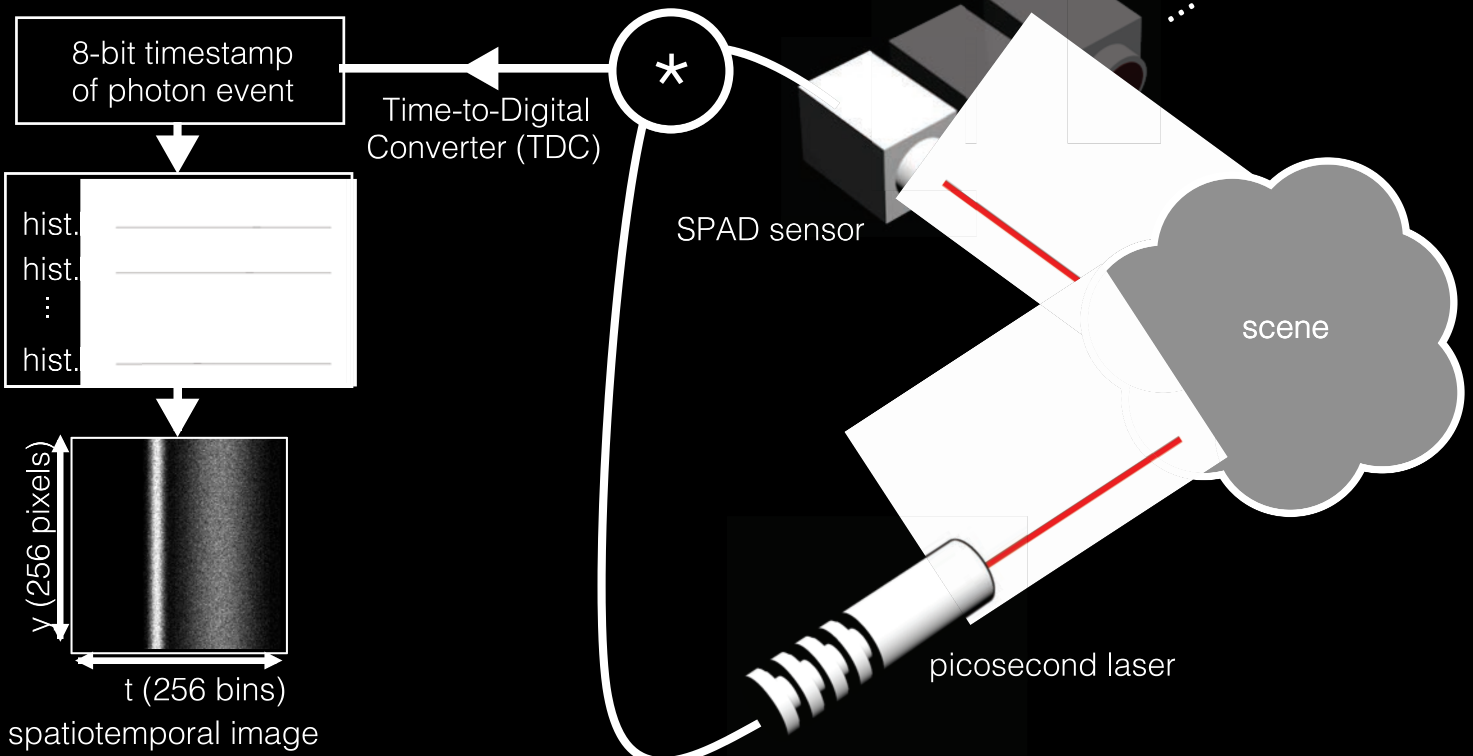
Challenges

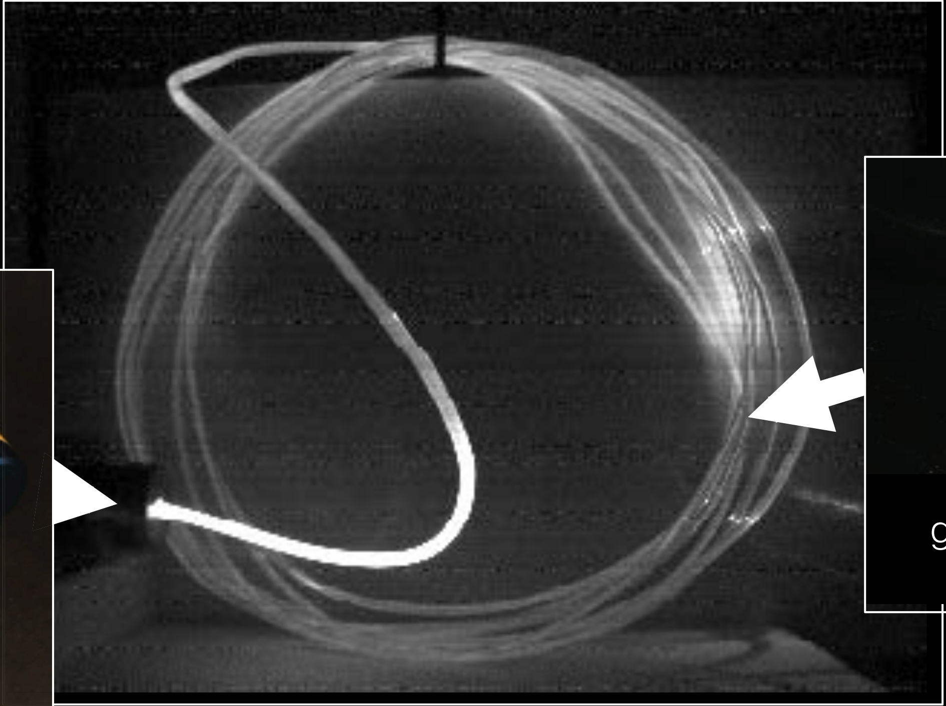
1. Light efficiency
2. High-speed time stamping

Single-photon Avalanche Diode (SPAD)

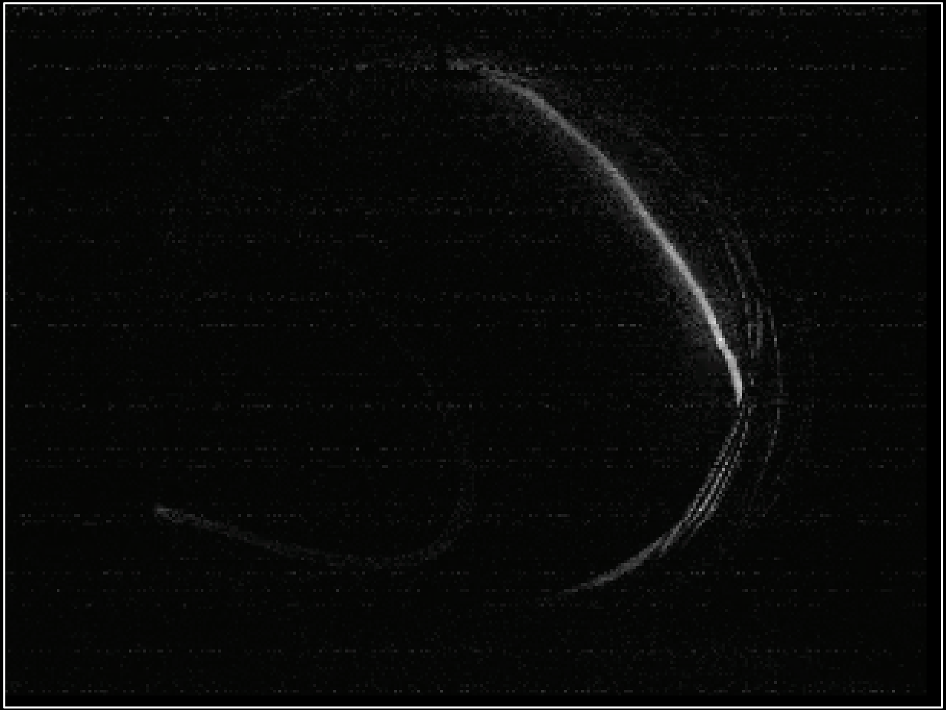


Single-photon Avalanche Diode Array



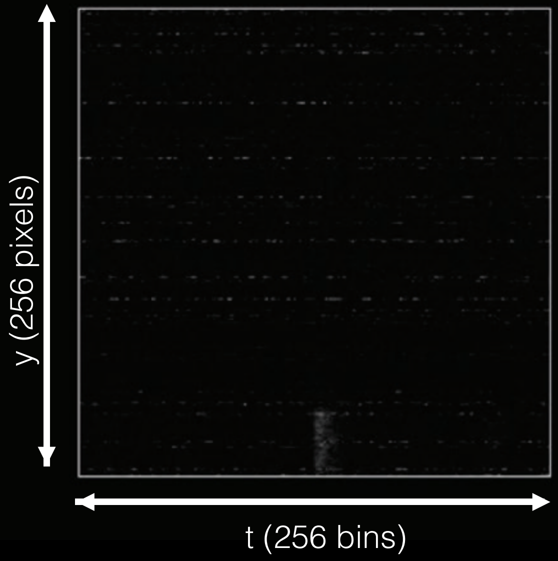


regular image

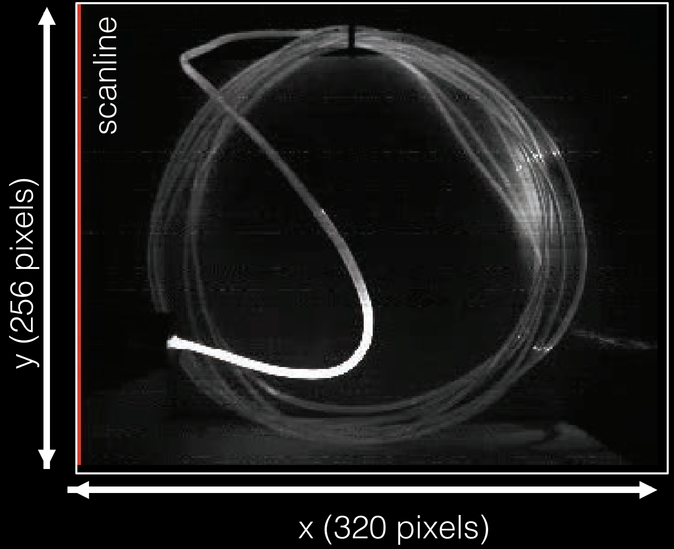


“transient” image

LinoSPAD Scanning Procedure

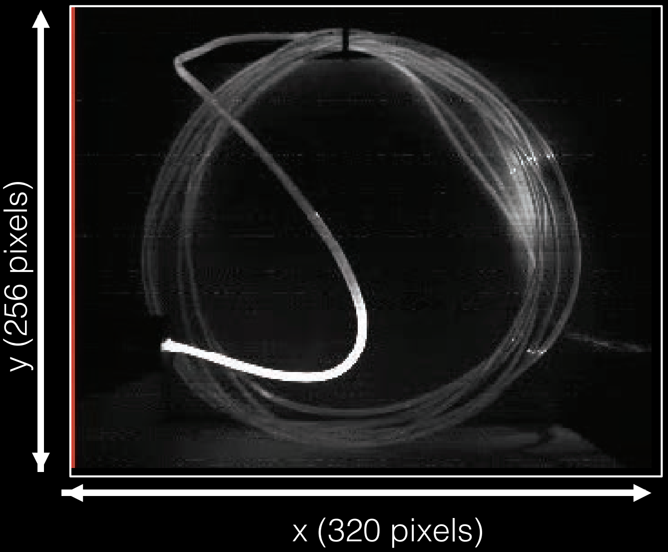


SPAD output



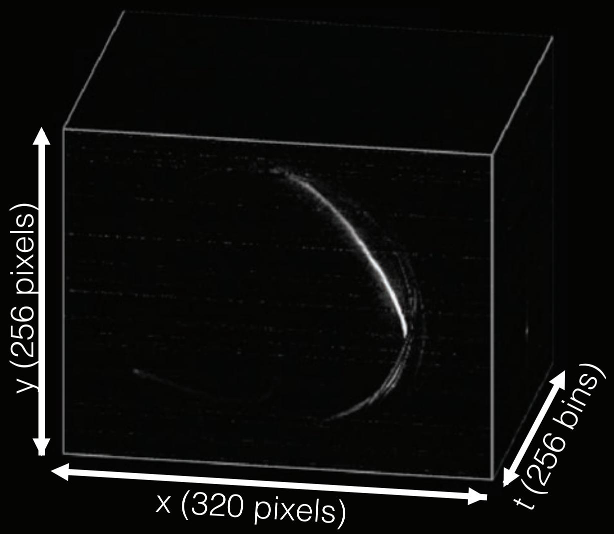
regular image

LinoSPAD Scanning Procedure

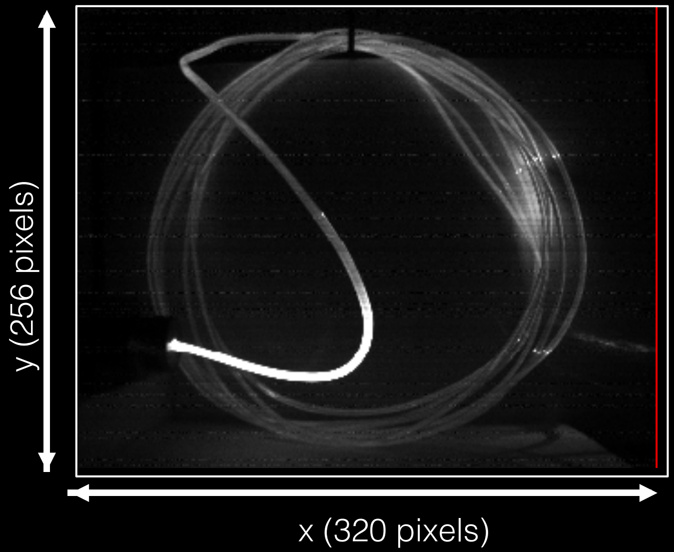


regular image

LinoSPAD Scanning Procedure



transient image



regular image

Reconstructing Transient Images

log likelihood of image formation model with Poisson noise

prior on 3D spatio-temporal data, e.g. 3DTV, NLM, BM3D, ...

$$\underset{\{\tau\}}{\text{minimize}} \quad -\log\left(p\left(\mathbf{h}|\mathbf{A}\boldsymbol{\tau}\right)\right) + \Gamma\left(\boldsymbol{\tau}\right)$$

$$\text{subject to} \quad 0 \leq \boldsymbol{\tau}$$

nonnegativity constraints

Splitting-based ADMM Solver:

$$\underset{\{\tau\}}{\text{minimize}} \quad -\log\left(p\left(\mathbf{h}|\mathbf{z}_1\right)\right) + \mathcal{I}_{\mathbb{R}_+}\left(\mathbf{z}_2\right) + \Gamma\left(\mathbf{z}_3\right)$$

$$\text{subject to} \quad \underbrace{\begin{bmatrix}\mathbf{A}\\ \mathbf{I}\\ \mathbf{I}\end{bmatrix}}_{\mathbf{K}} \boldsymbol{\tau} - \underbrace{\begin{bmatrix}\mathbf{z}_1\\ \mathbf{z}_2\\ \mathbf{z}_3\end{bmatrix}}_{\mathbf{z}} = 0$$

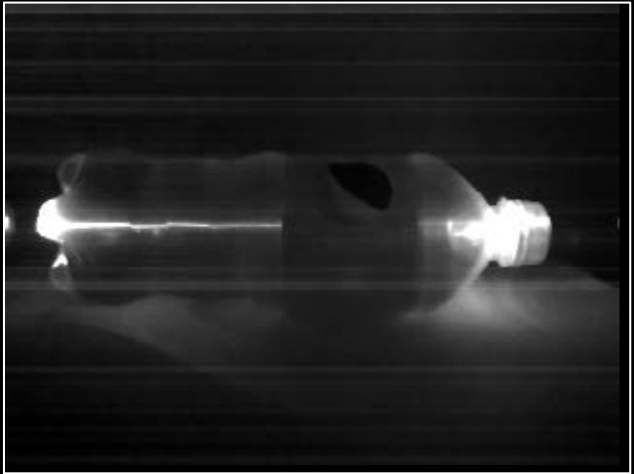
Algorithm 1 ADMM-based denoising and deconvolution

- **for** $k = 1$ **to** M
  - 3D deconvolution, closed-form solution → $\boldsymbol{\tau} \leftarrow \underset{\{\boldsymbol{\tau}\}}{\arg\min} \frac{1}{2}\left\|\mathbf{K}\boldsymbol{\tau} - \mathbf{z} + \mathbf{u}\right\|_2^2$
  - Poisson denoising, closed-form solution → $\mathbf{z}_1 \leftarrow \underset{\{\mathbf{z}_1\}}{\arg\min} -\log\left(p\left(\mathbf{h}|\mathbf{z}_1\right)\right) + \frac{\rho}{2}\left\|\mathbf{A}\boldsymbol{\tau} + \mathbf{u}_1 - \mathbf{z}_1\right\|_2^2$
  - projection, closed-form solution → $\mathbf{z}_2 \leftarrow \underset{\{\mathbf{z}_2\}}{\arg\min} \mathcal{I}_{\mathbb{R}_+}\left(\mathbf{z}_2\right) + \frac{\rho}{2}\left\|\boldsymbol{\tau} + \mathbf{u}_2 - \mathbf{z}_2\right\|_2^2$
  - proximal operator of prior, closed-form solution → $\mathbf{z}_3 \leftarrow \underset{\{\mathbf{z}_3\}}{\arg\min} \Gamma\left(\mathbf{z}_3\right) + \frac{\rho}{2}\left\|\boldsymbol{\tau} + \mathbf{u}_3 - \mathbf{z}_3\right\|_2^2$
  - sum of terms, closed-form solution → $\mathbf{u} \leftarrow \mathbf{u} + \mathbf{K}\boldsymbol{\tau} - \mathbf{z}$
- **end for**

O'Toole et al., CVPR 2017

Transient Imaging Results

regular image


regular image

O'Toole et al., CVPR 2017

Transient Imaging Results

regular image

transient image

O’Toole et al., CVPR 2017

Transient Imaging Results

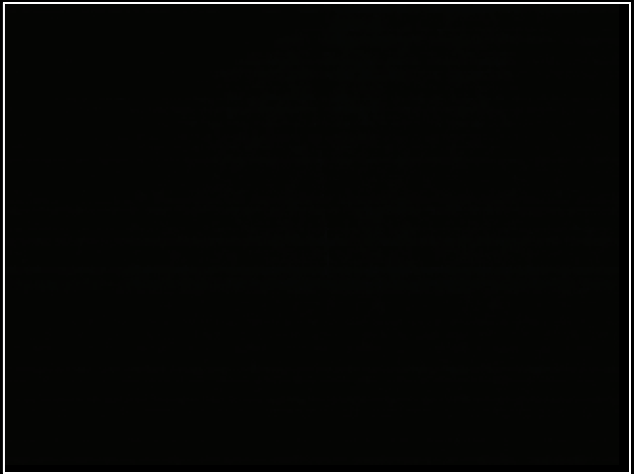
regular image

regular image

O’Toole et al., CVPR 2017

Transient Imaging Results

regular image



transient image

O'Toole et al., CVPR 2017

Pushing the Limits of Transient Imaging

Acquisition at 25 Hz with 64x80 resolution

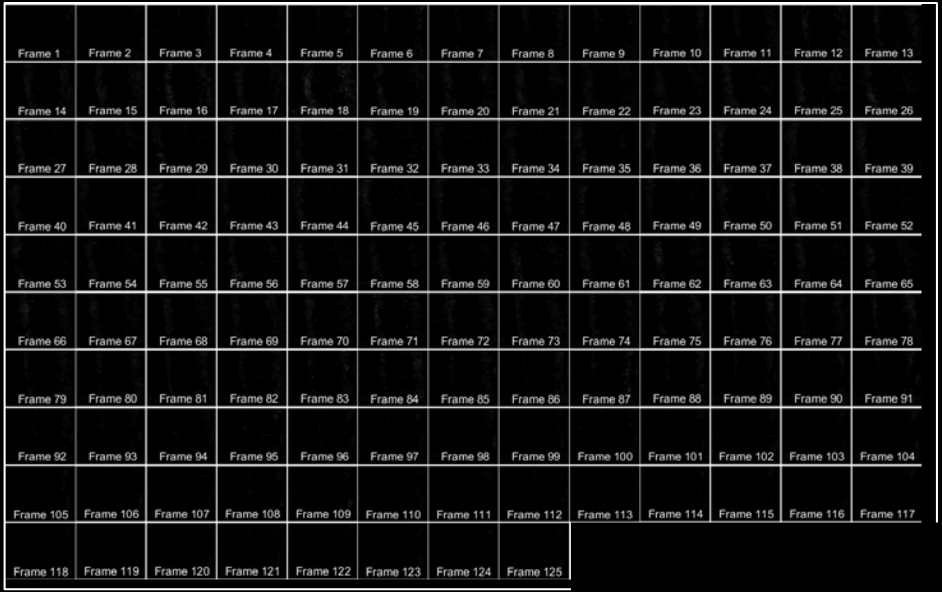
scene under laser light

transient image
(RAW)

transient image
(processed)

Lindell et al., ICCP 2018

scene under laser light



transient video (processed)

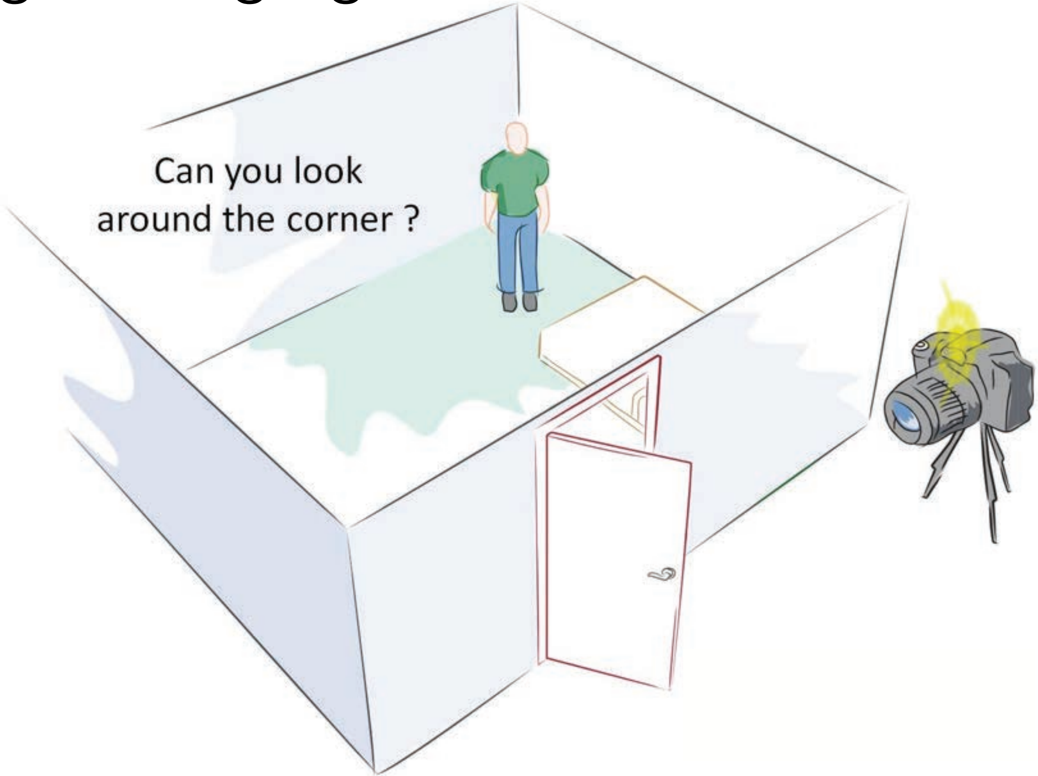
Lindell et al., ICCP 2018

Applications of Transient Imaging

- depth estimation
- direct-global illumination separation
- light transport analysis
- fundamentally new imaging modality that could enable new capabilities for image processing & computer vision algorithms … ongoing work

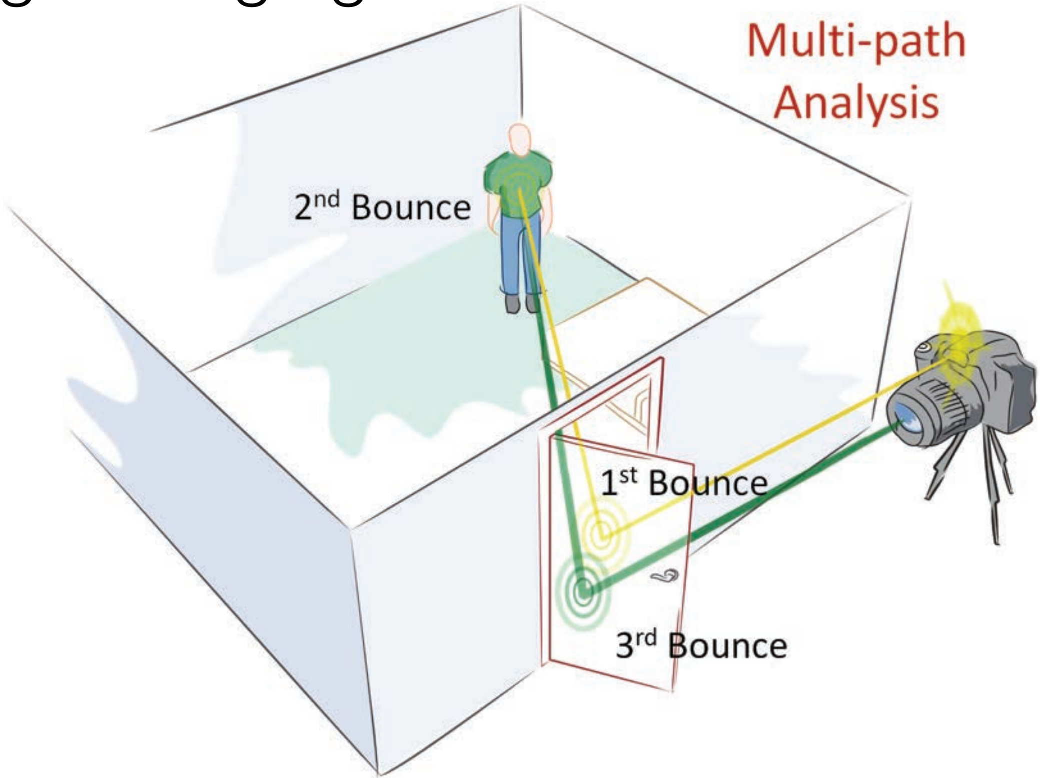
- enables *Non-line-of-sight (NLOS) Imaging*

Non-line-of-sight Imaging

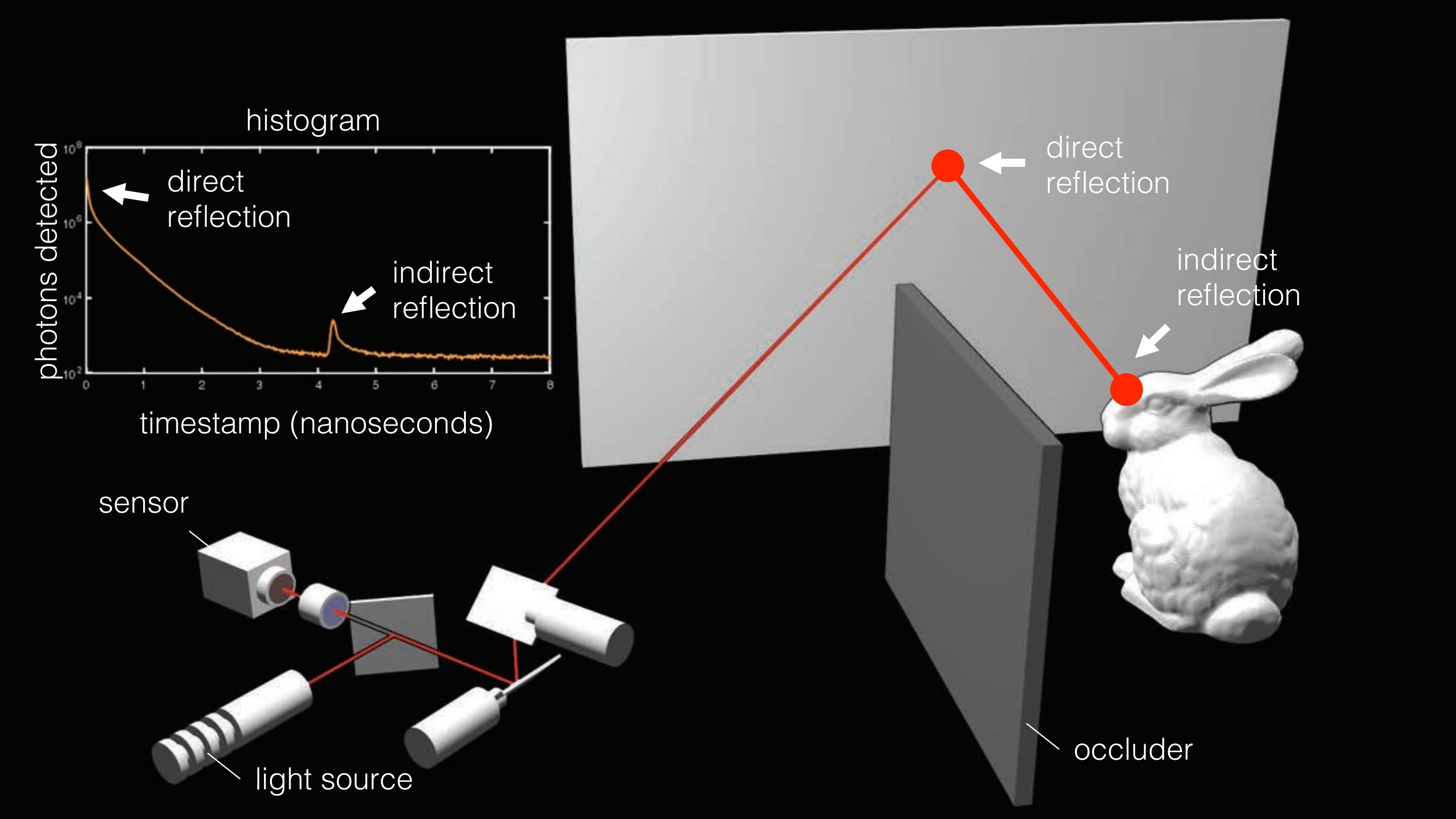


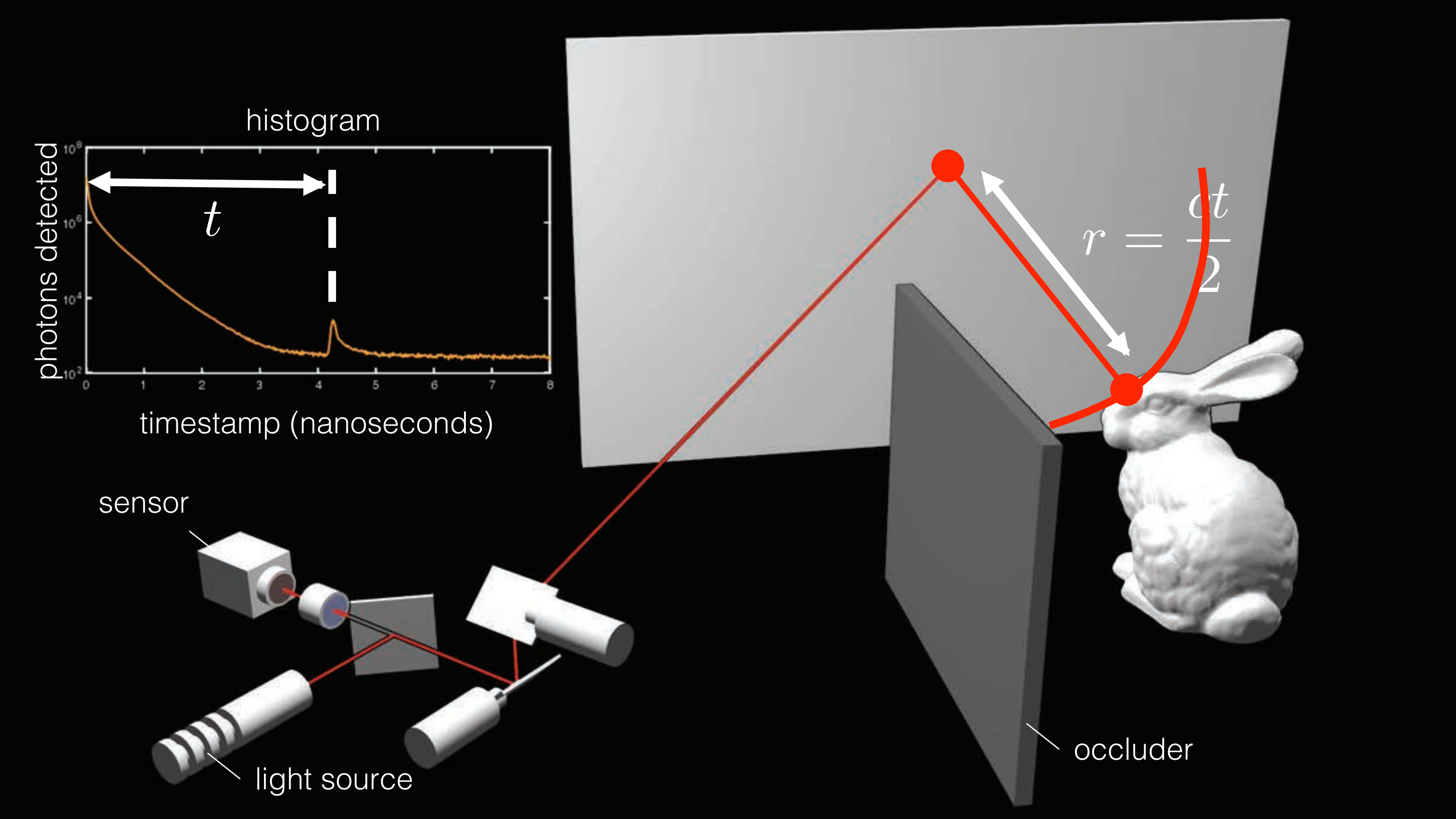
Raskar et al. 2012

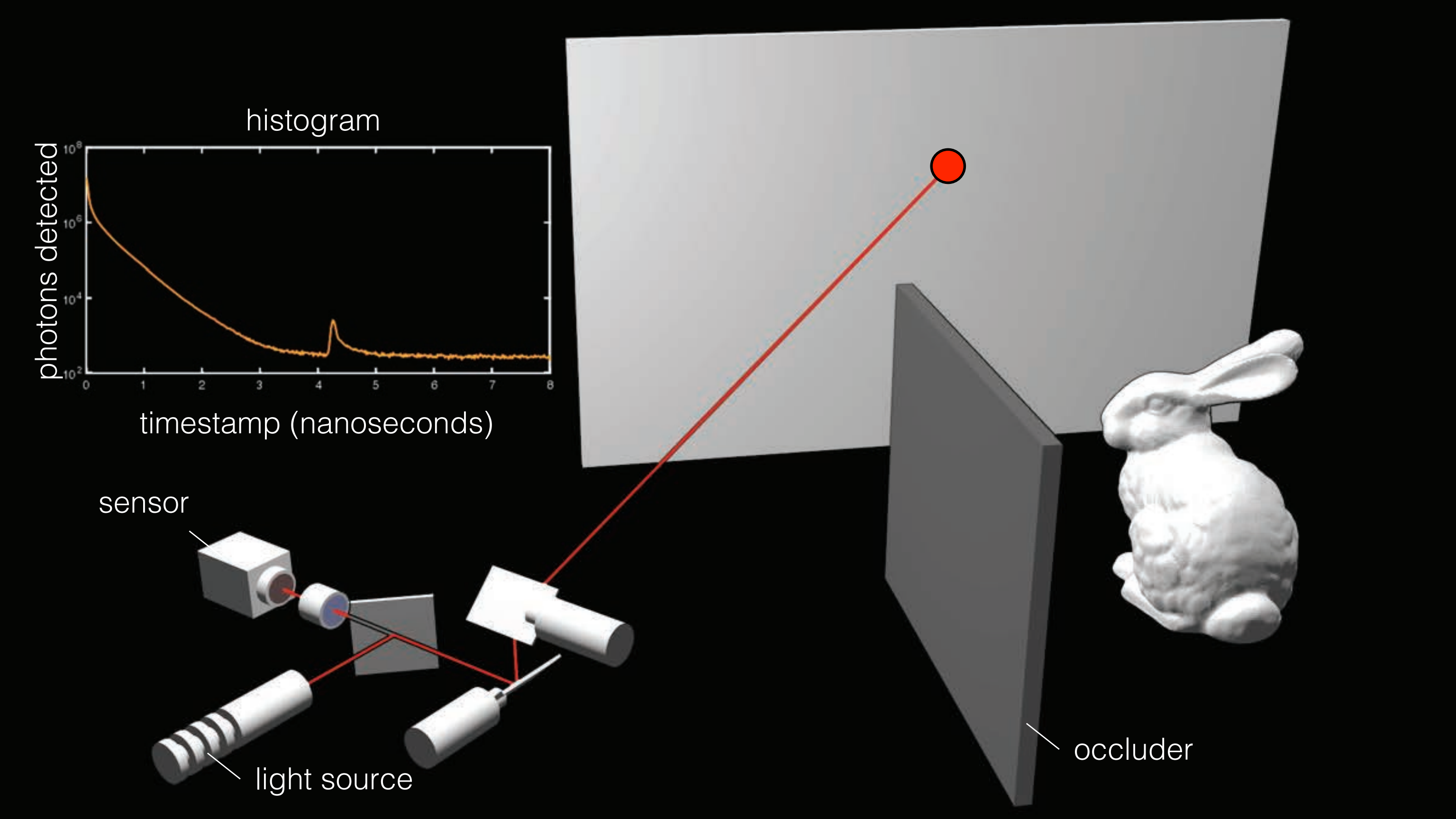
Non-line-of-sight Imaging

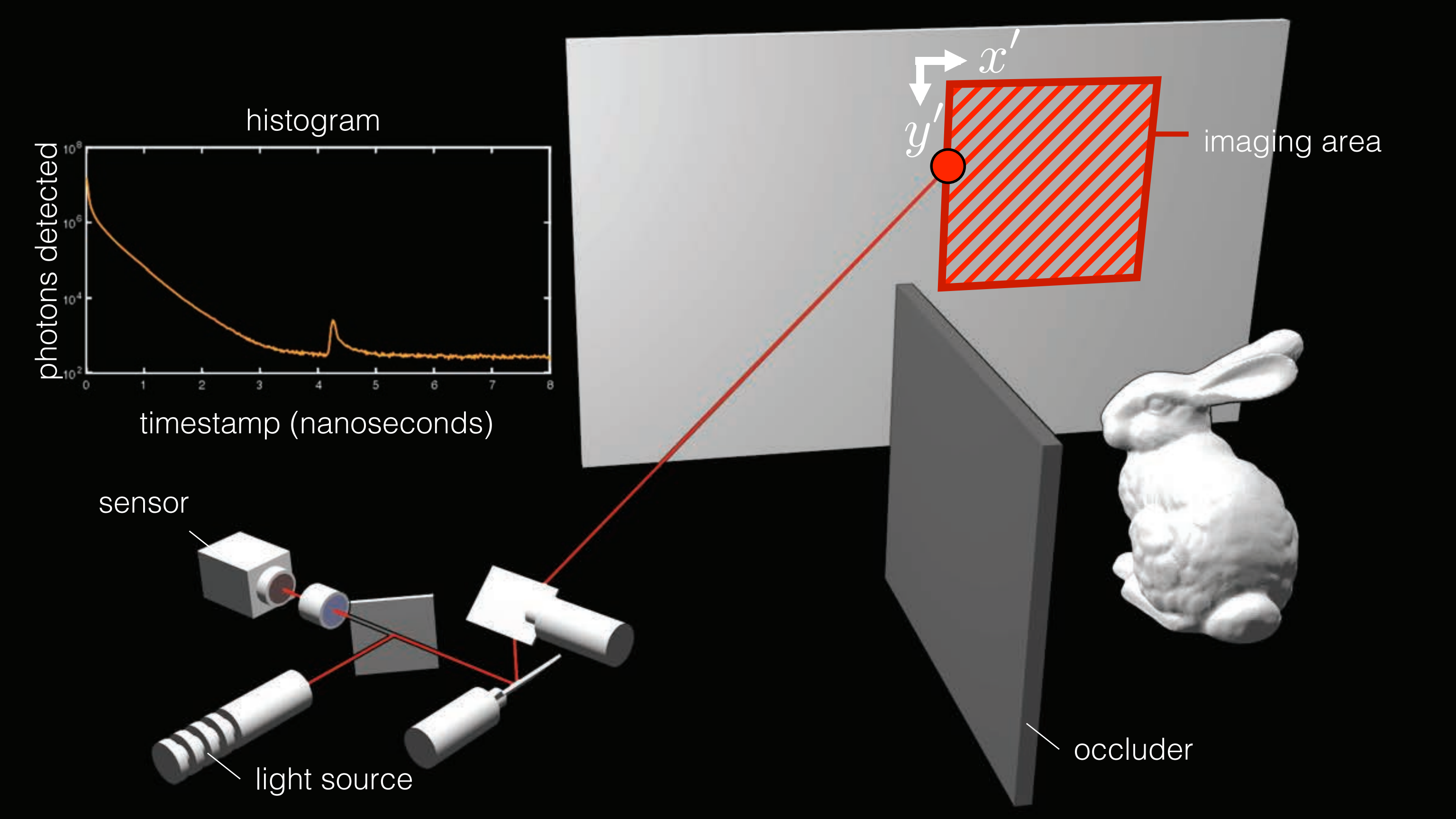


Raskar et al. 2012

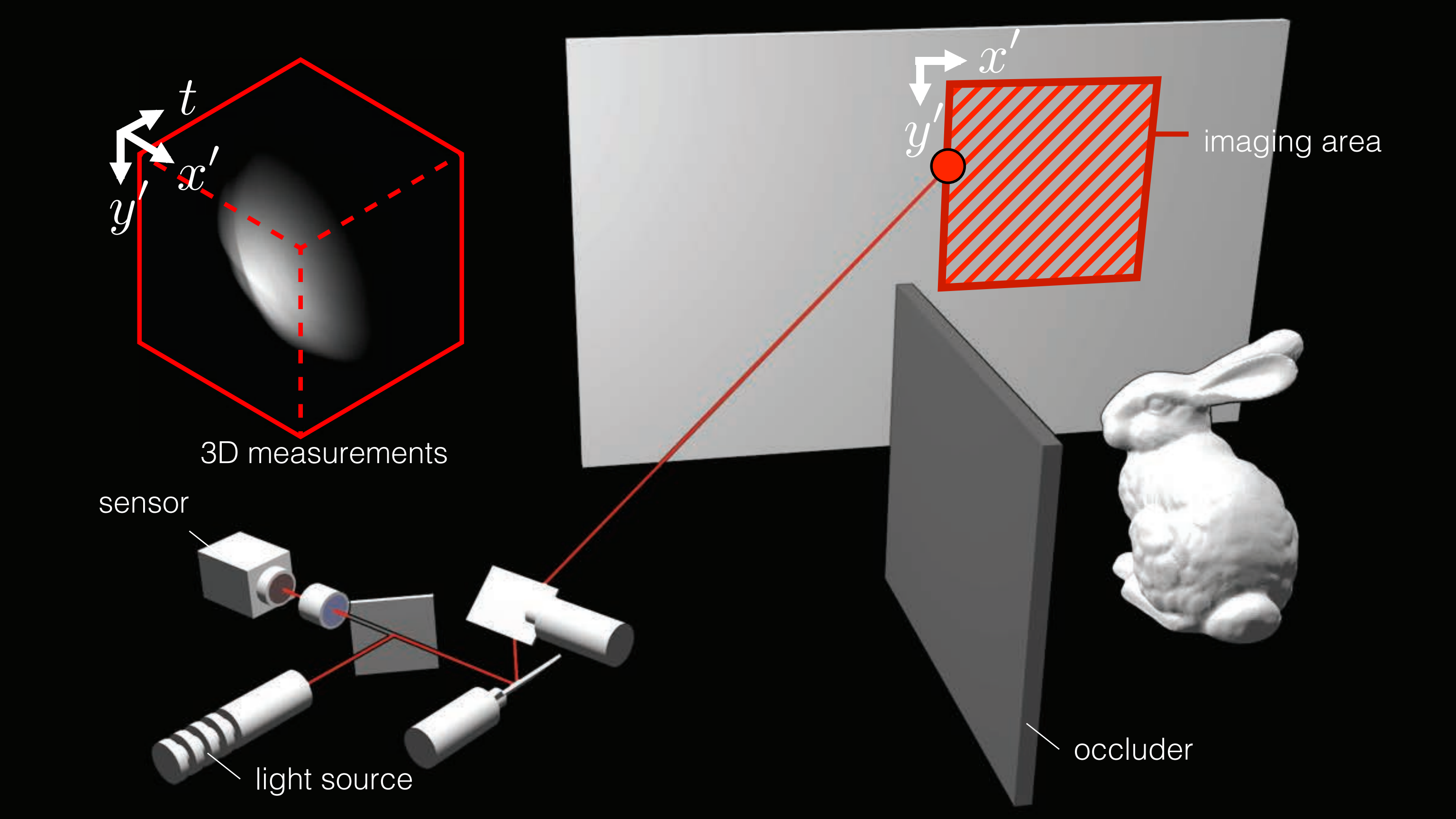

histogram
photons detected
direct
reflection
indirect
reflection
timestamp (nanoseconds)
0
1
2
3
4
5
6
7
8
direct
reflection
indirect
reflection
sensor
light source
occluder


histogram
photons detected
10^8
10^6
10^4
10^2
0
1
2
3
4
5
6
7
8
t
timestamp (nanoseconds)
$r = \frac{ct}{2}$
sensor
light source
occluder

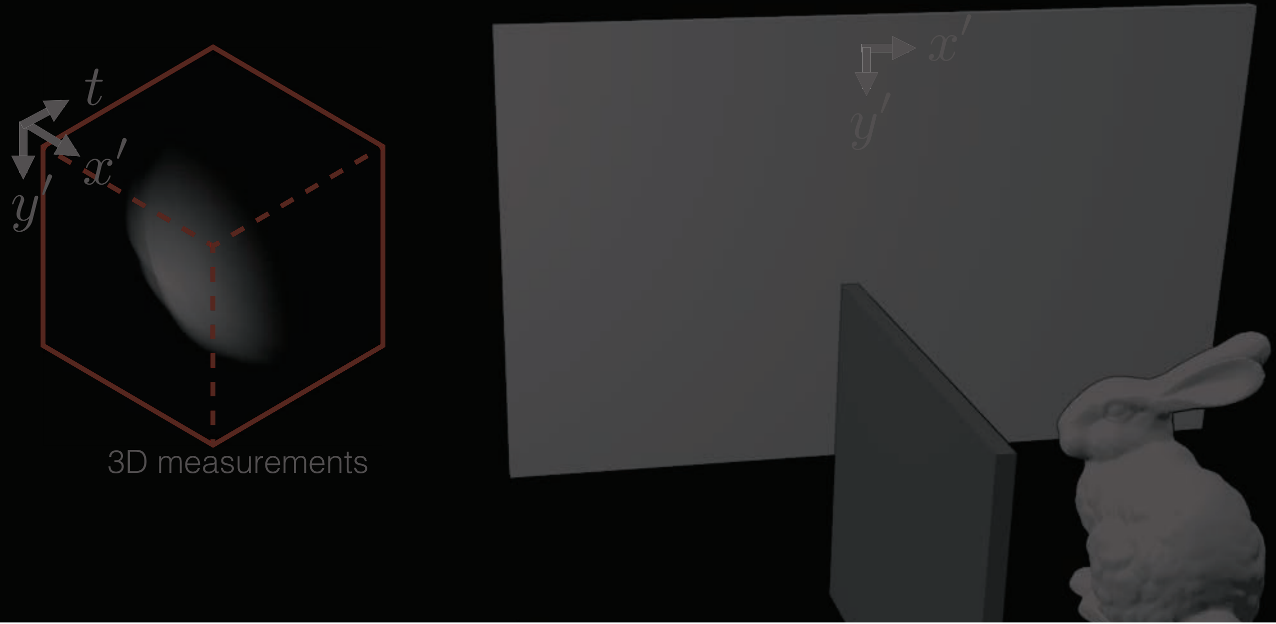

histogram
photons detected
timestamp (nanoseconds)
sensor
light source
occluder



histogram
photons detected
timestamp (nanoseconds)
x'
y'
imaging area
sensor
light source
occluder

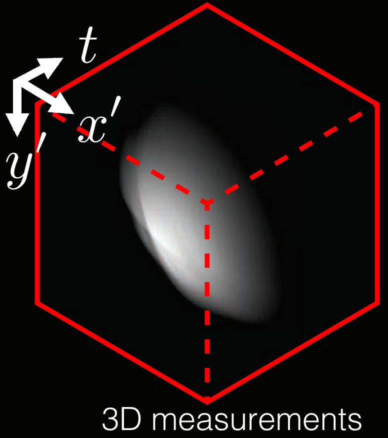


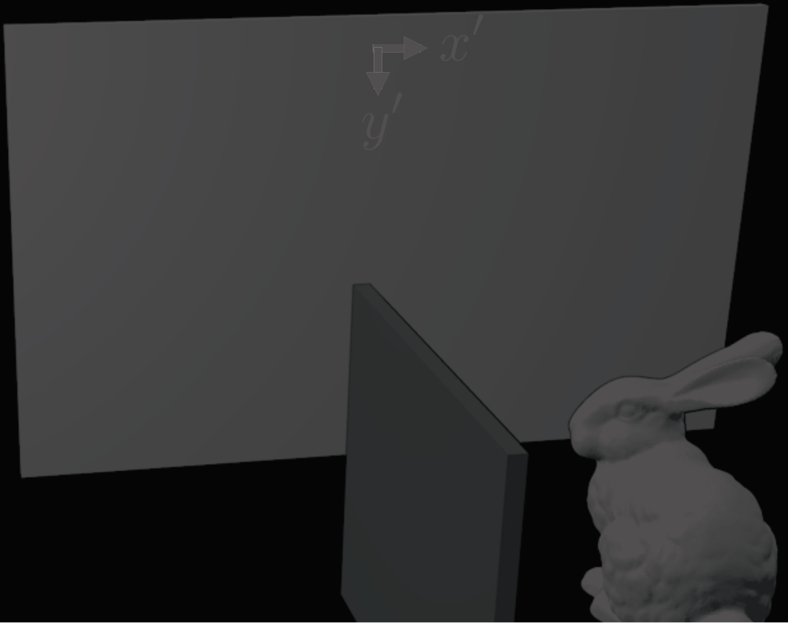
t
x'
y'
3D measurements
imaging area
sensor
light source
occluder



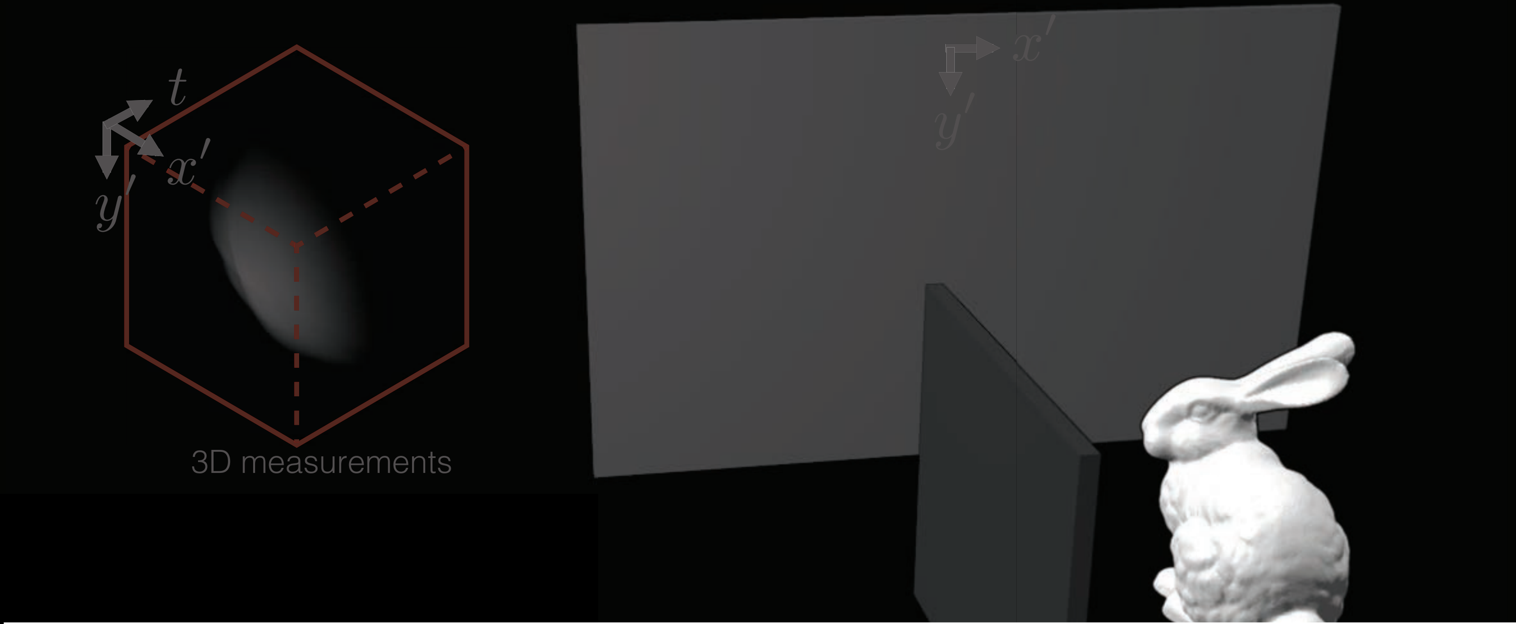
$$\tau(x',y',t) = \iiint_{\Omega} \frac{1}{r_l^2 r^2} \delta\left(r_l + r - tc\right) \cdot \rho(x,y,z)\ dx\ dy\ dz$$

3D measurements — radiometric term — geometric term — hidden 3D volume



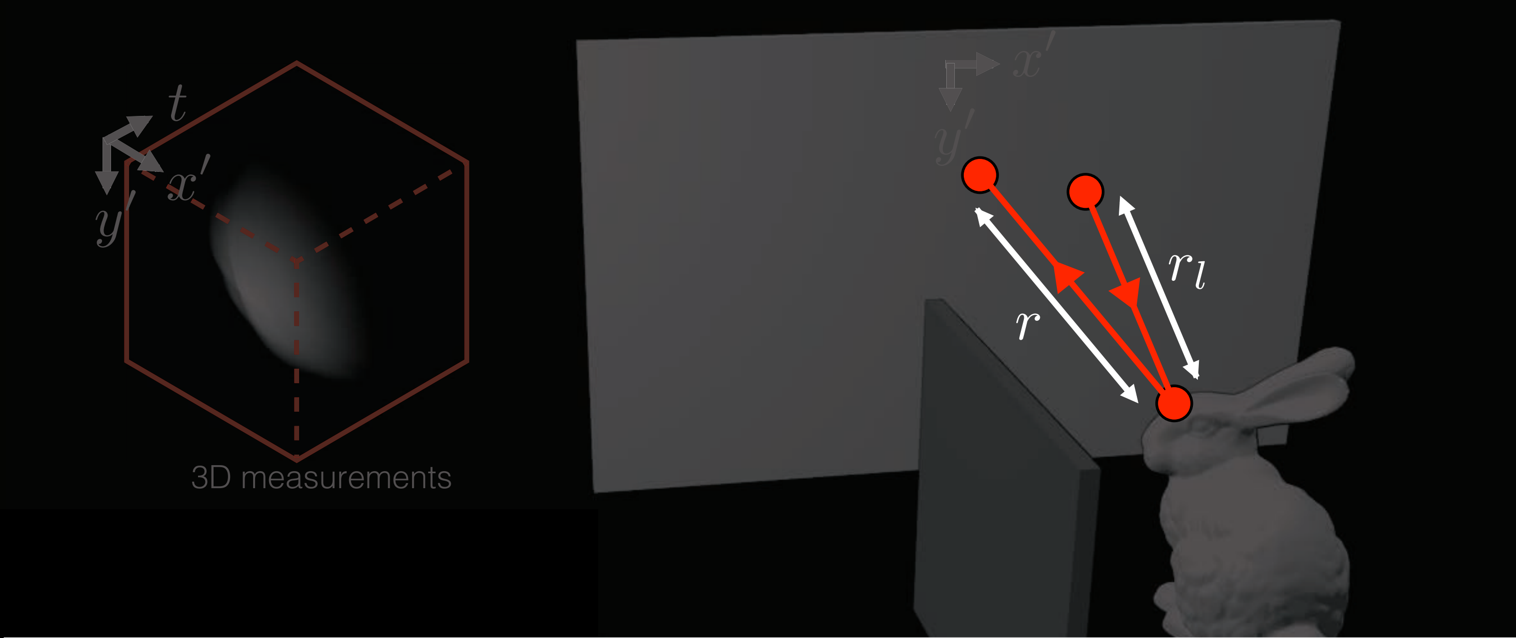


$$\underbrace{\tau(x',y',t)}_{\text{3D measurements}} = \iiint_\Omega \underbrace{\frac{1}{r_l^2 r^2}}_{\text{radiometric term}} \underbrace{\delta\left(r_l + r - tc\right)}_{\text{geometric term}} \cdot \underbrace{\rho(x,y,z)}_{\text{hidden 3D volume}}\, dx\, dy\, dz$$



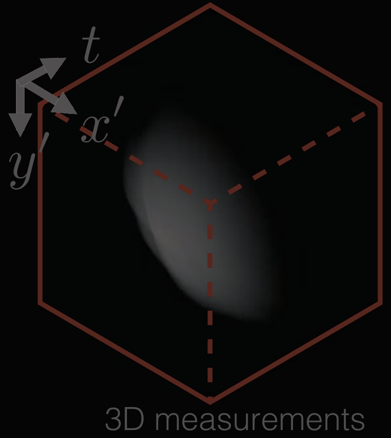
$$\tau(x', y', t) = \iiint_\Omega \frac{1}{r_l^2 r^2} \delta\left(r_l + r - tc\right) \cdot \rho(x, y, z)\, dx\ dy\ dz$$

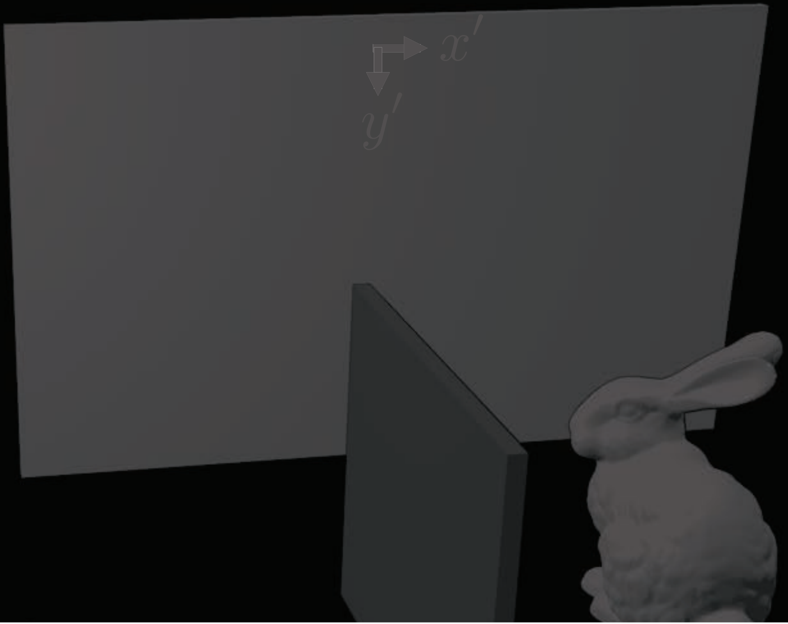
3D measurements — radiometric term — geometric term — hidden 3D volume



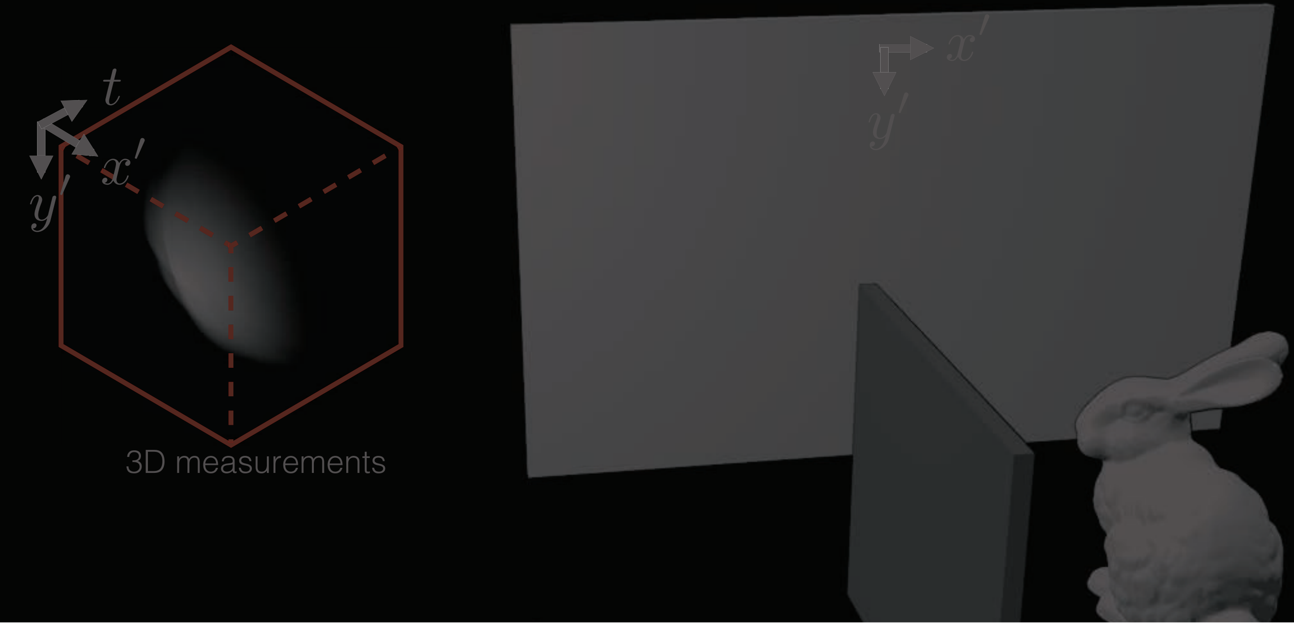
$$\tau(x',y',t) = \iiint_\Omega \frac{1}{r_l^2 r^2}\,\delta\left(r_l + r - tc\right) \cdot \rho(x,y,z)\; dx\; dy\; dz$$

3D measurements — radiometric term — geometric term — hidden 3D volume





$$\tau(x', y', t) = \iiint_{\Omega} \frac{1}{r_l^2 r^2} \delta\left(r_l + r - tc\right) \cdot \rho(x, y, z)\ dx\ dy\ dz$$



$$\tau(x', y', t) = \iiint_\Omega \frac{1}{r_l^2 r^2} \delta\left(r_l + r - tc\right) \cdot \rho(x, y, z)\ dx\ dy\ dz$$

$$\underbrace{\qquad}_{\boldsymbol{\tau}} = \underbrace{\qquad}_{\mathbf{A}} \times \underbrace{\qquad}_{\boldsymbol{\rho}}$$

NLOS image formation mode:

$$\tau = \mathbf{A}\rho$$

measurements $n^3 \times 1$

transport matrix $n^3 \times n^3$

unknown volume $n^3 \times 1$

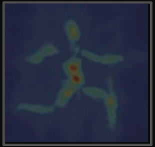
PROBLEM: A extremely large in practice

- for n=100, A has 1 trillion elements
- for n=1000, sparse A needs 9 petabyte of memory

Backpropagation [Velten 12, Buttafava 15]

Flops: $O(n^5)$

Memory: $O(n^3)$



Iterative Inversion [Gupta 12, Wu 12, Heide 13]

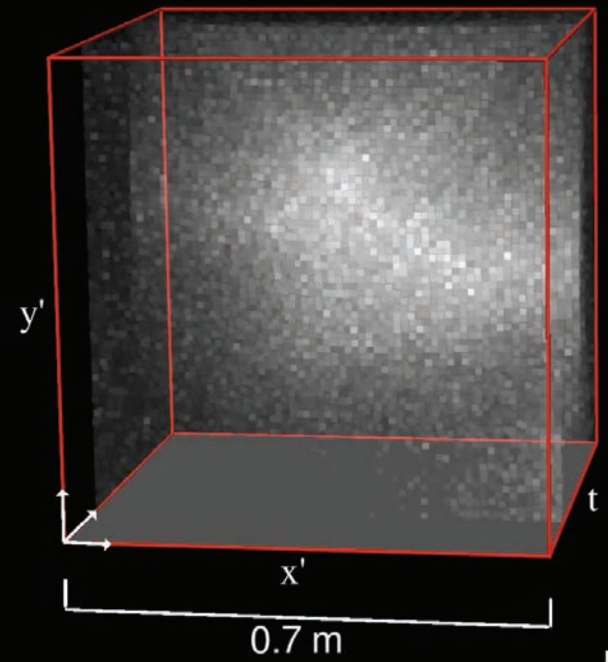
Flops: $O(n^5)$ per iter.

Memory: $O(n^5)$

Challenges of NLOS Imaging

1. Light efficiency, high-speed time stamping
2. Efficient Scanning
3. Large-scale inverse problem
4. Accurate (and invertible) NLOS light transport model

Confocal Non-line-of-sight Imaging



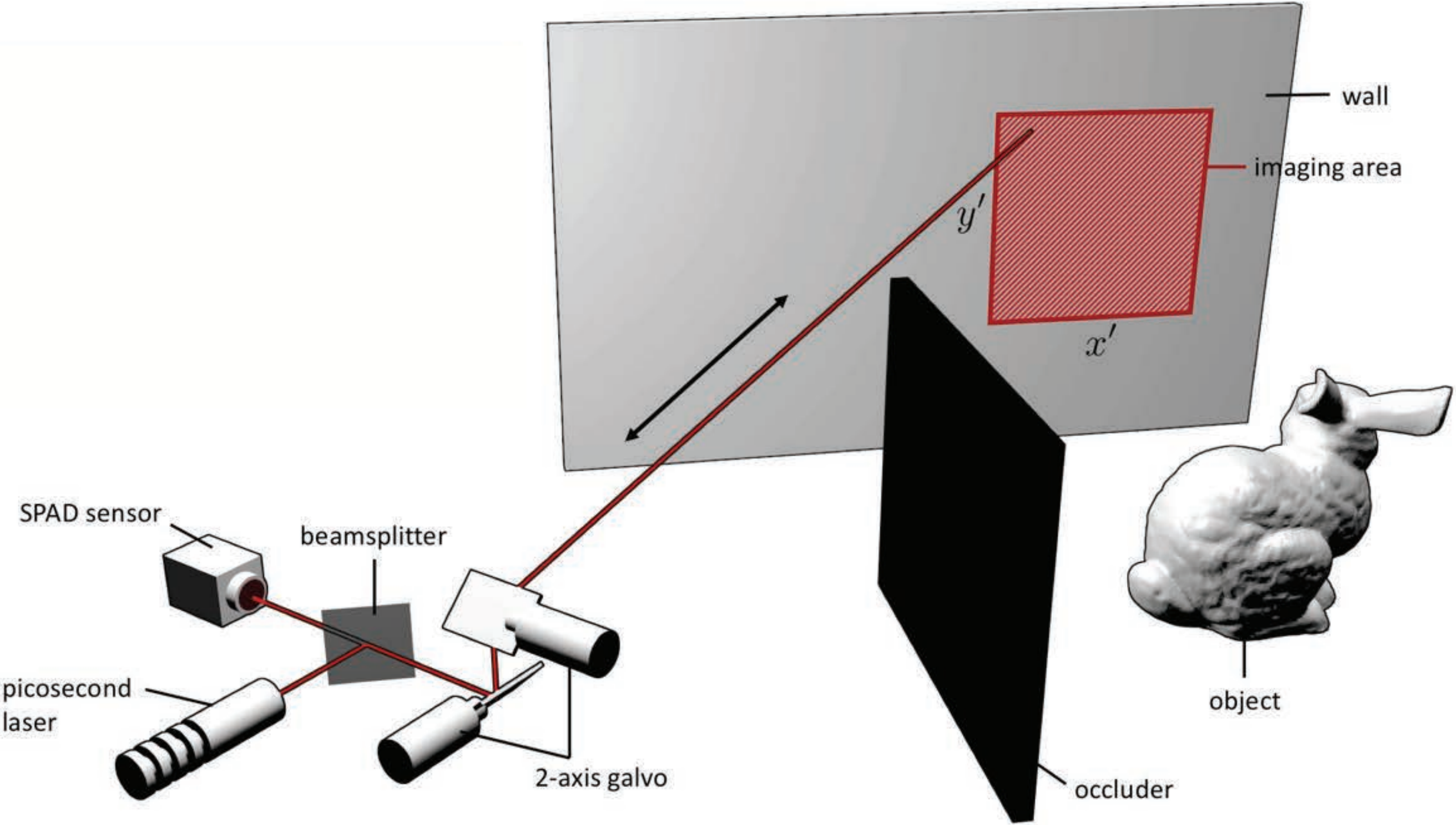
Maximum Intensity Projection

Challenges of NLOS Imaging

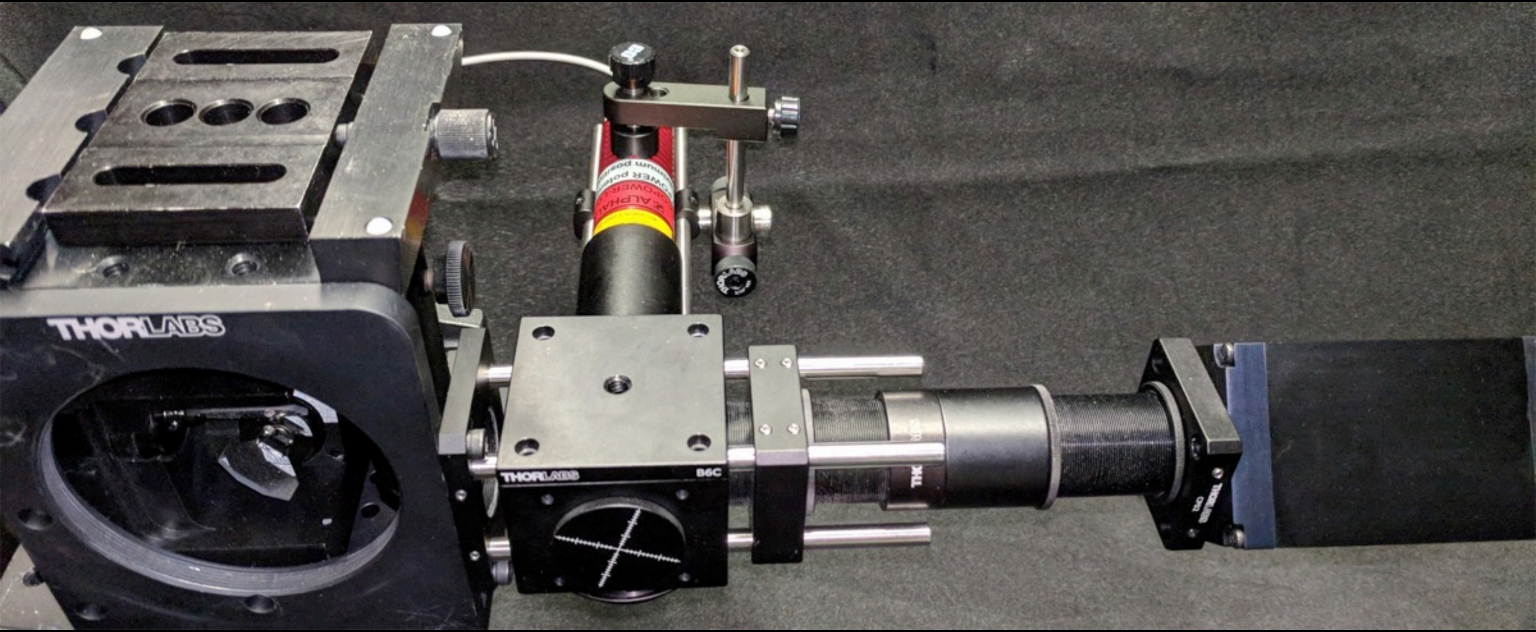
1. Light efficiency, high-speed time stamping → SPADs
   - compatible with LIDAR systems
   - low-cost fabrication; silicon & CMOS
   - increasing availability of detectors
2. Efficient Scanning
3. Large-scale inverse problem
4. Accurate (and invertible) NLOS light transport model

Challenges of NLOS Imaging

1. Light efficiency, high-speed time stamping
2. Efficient Scanning
3. Large-scale inverse problem
4. Accurate (and invertible) NLOS light transport model

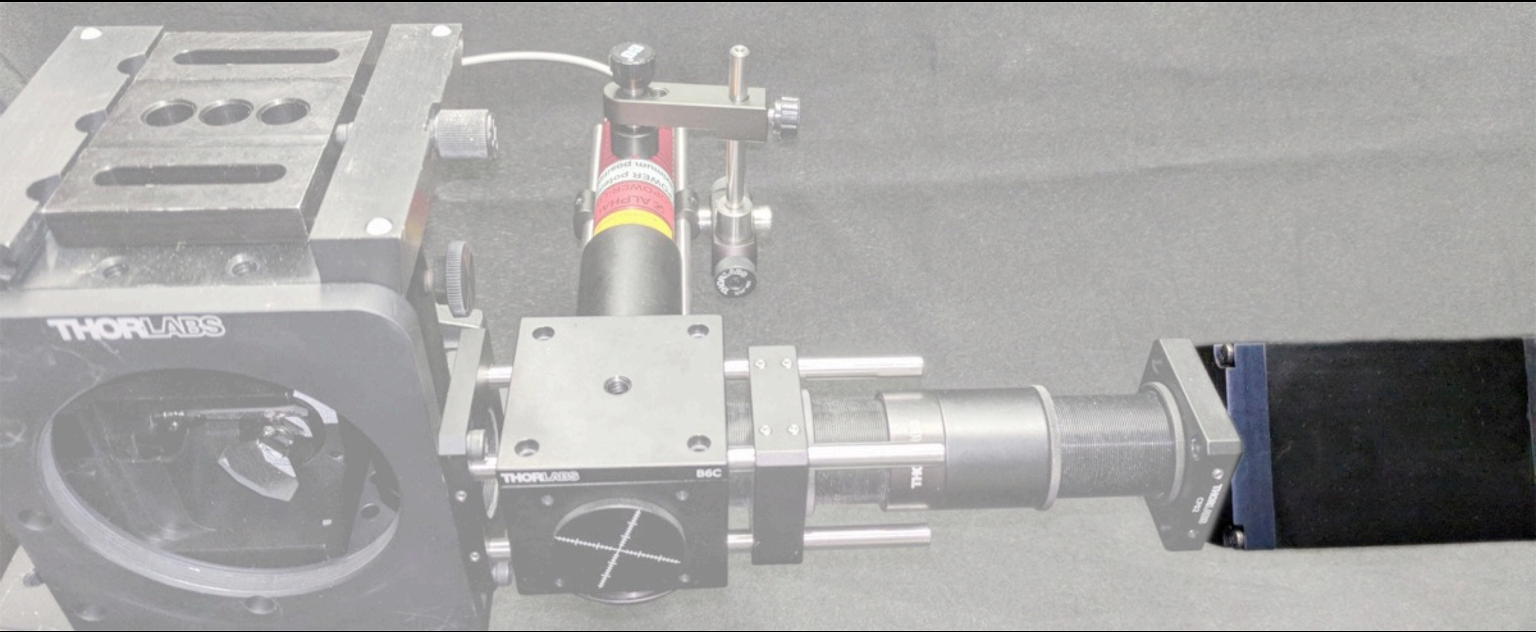

wall
imaging area
y'
x'
SPAD sensor
beamsplitter
picosecond laser
2-axis galvo
occluder
object

Confocal SPAD-based Scanning Setup



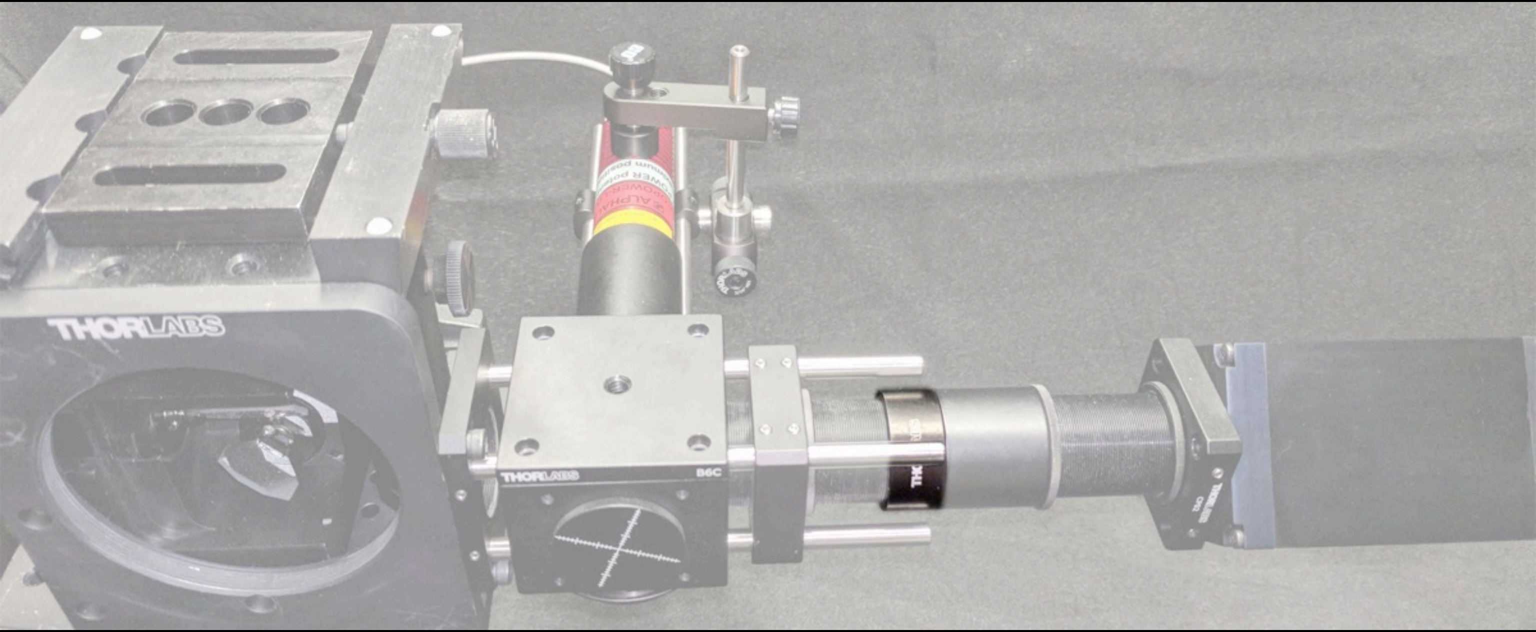
O'Toole et al., Nature 2018

Single Photon Avalanche Detector



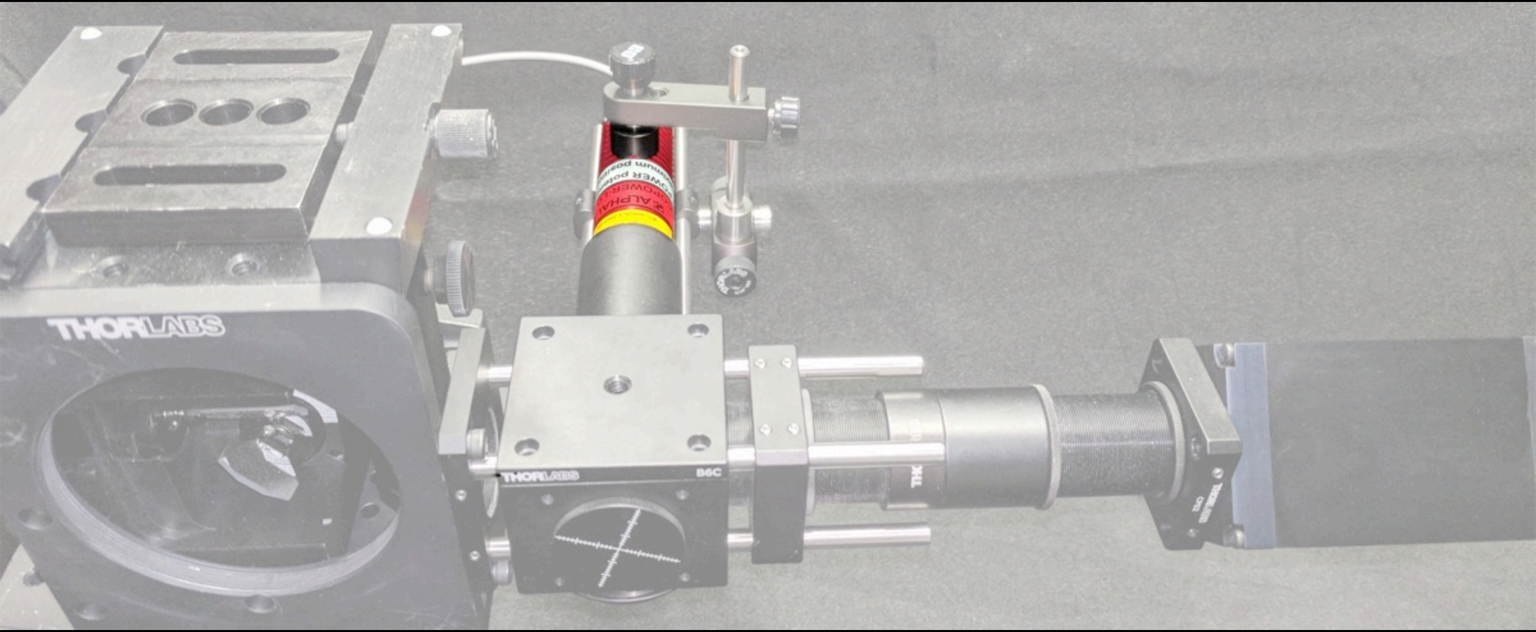
O'Toole et al., Nature 2018

Focusing Lens



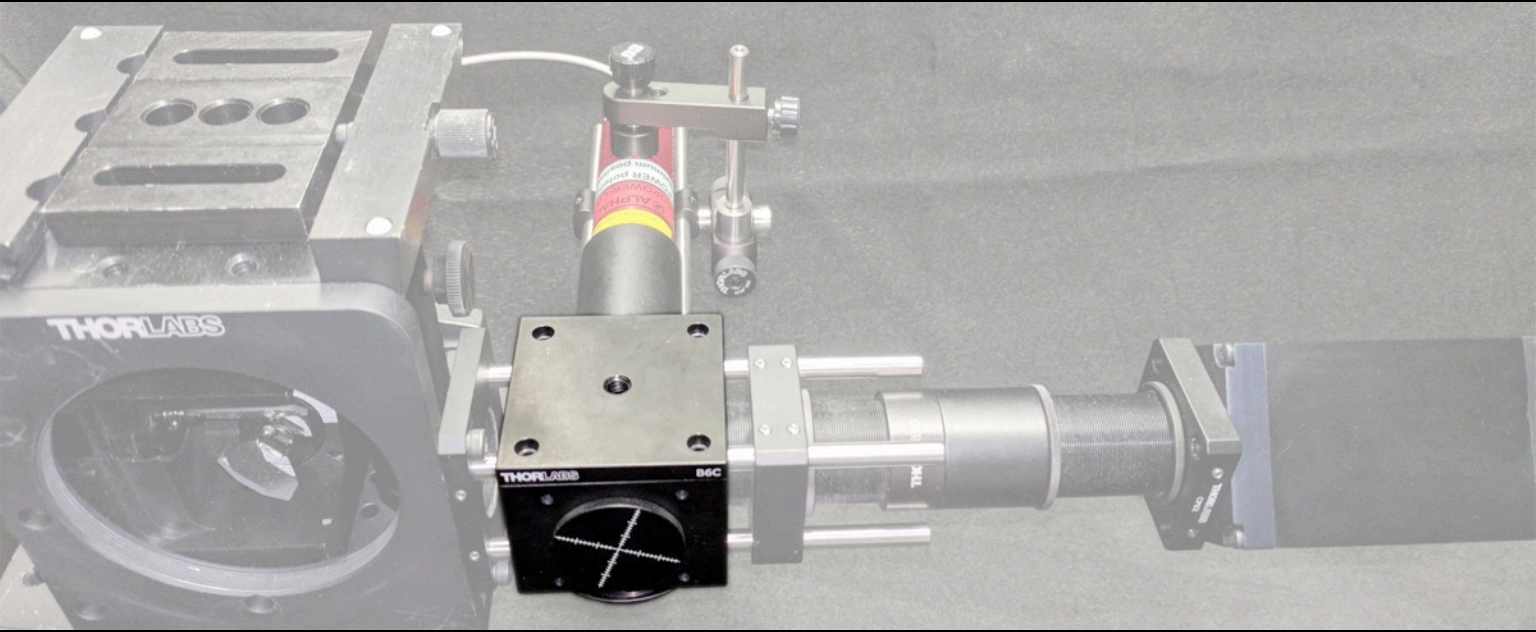
O'Toole et al., Nature 2018

Short-Pulsed Laser Illumination (50ps FWHM)



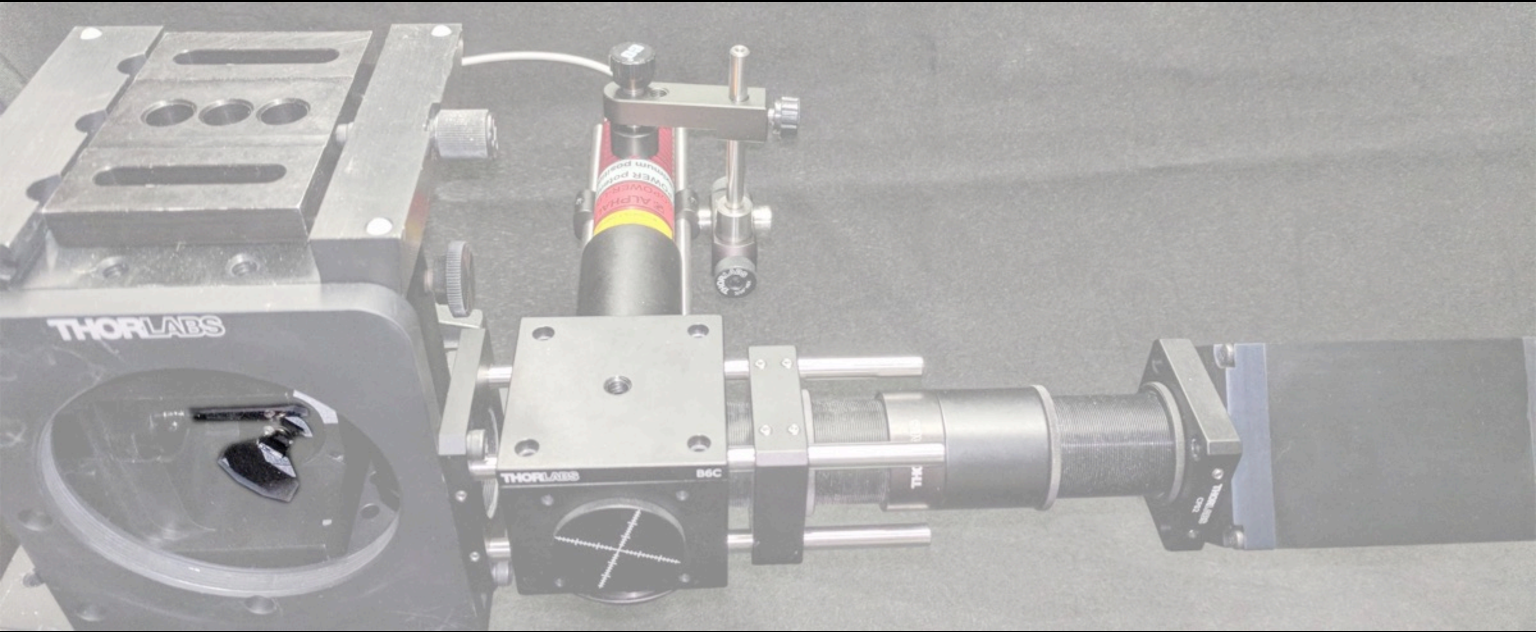
O’Toole et al., Nature 2018

Beam-Splitter in Coaxial Alignment



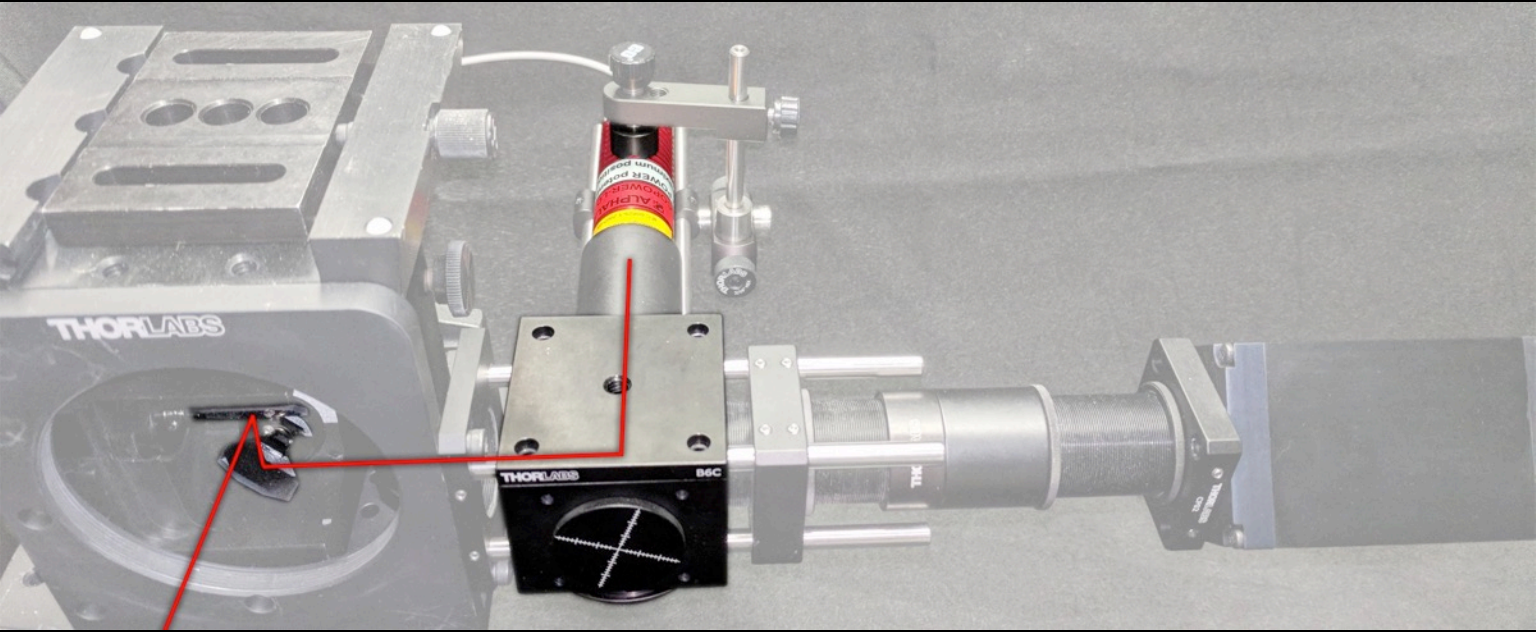
O'Toole et al., Nature 2018

Scanning Galvo Mirror System



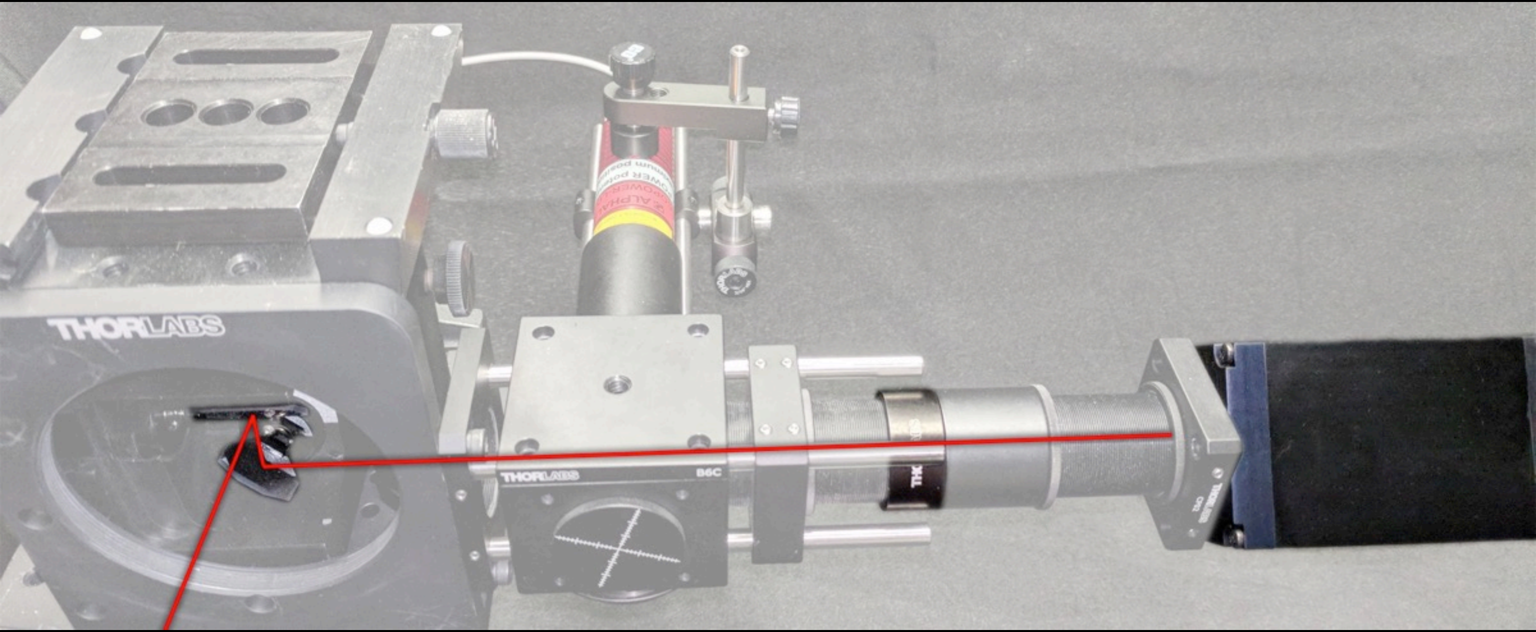
O’Toole et al., Nature 2018

Illumination Path



O’Toole et al., Nature 2018

Detection Path

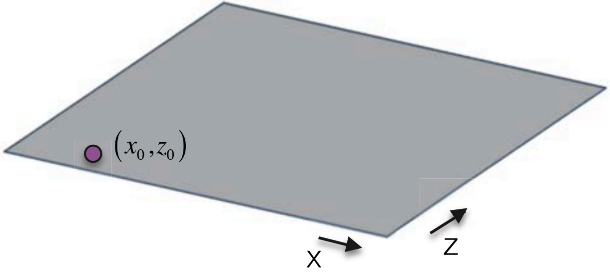


O'Toole et al., Nature 2018

Challenges of NLOS Imaging

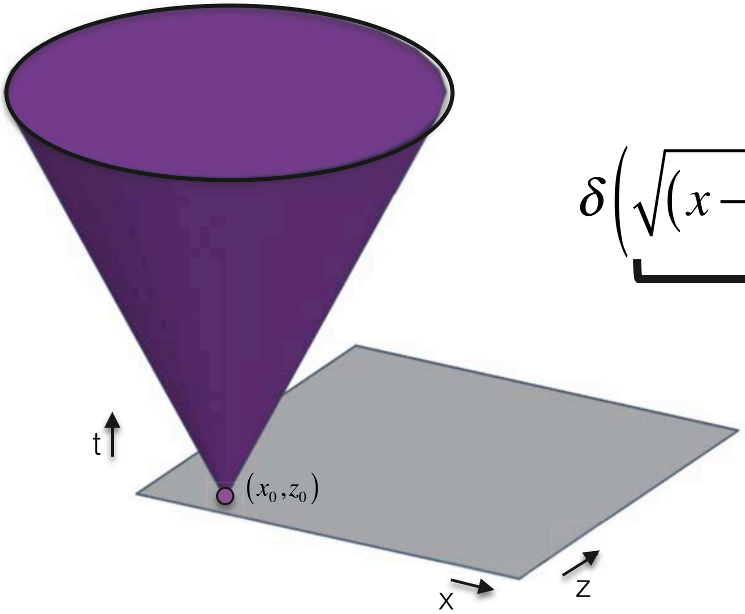
1. Light efficiency, high-speed time stamping
2. Efficient Scanning
3. Large-scale inverse problem
4. Accurate (and invertible) NLOS light transport model

Transient Light Transport Analysis



2D space

Transient Light Transport Analysis

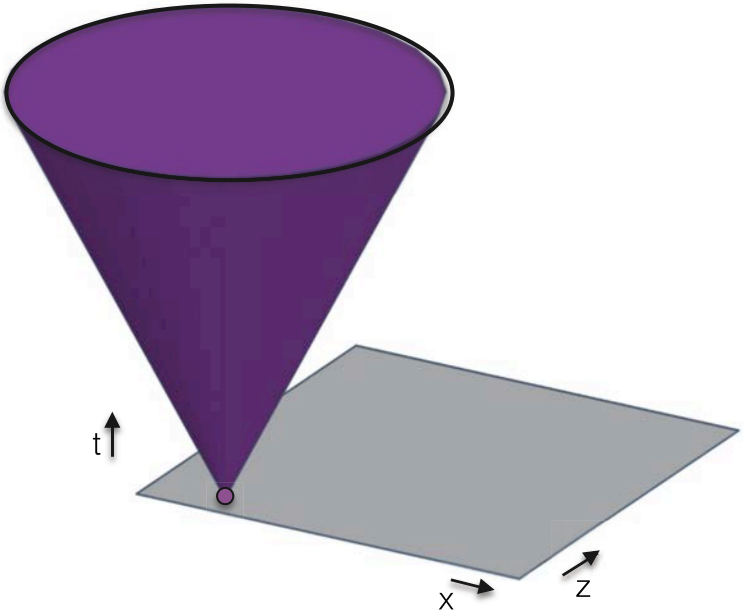


3D space-time

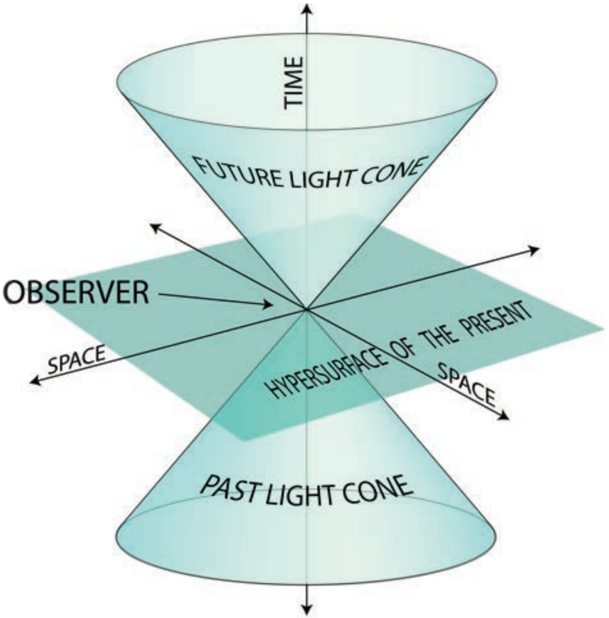
surface of hypercone

$$\delta\Big(\underbrace{\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}}_{r}-tc\Big)$$

Transient Light Transport Analysis

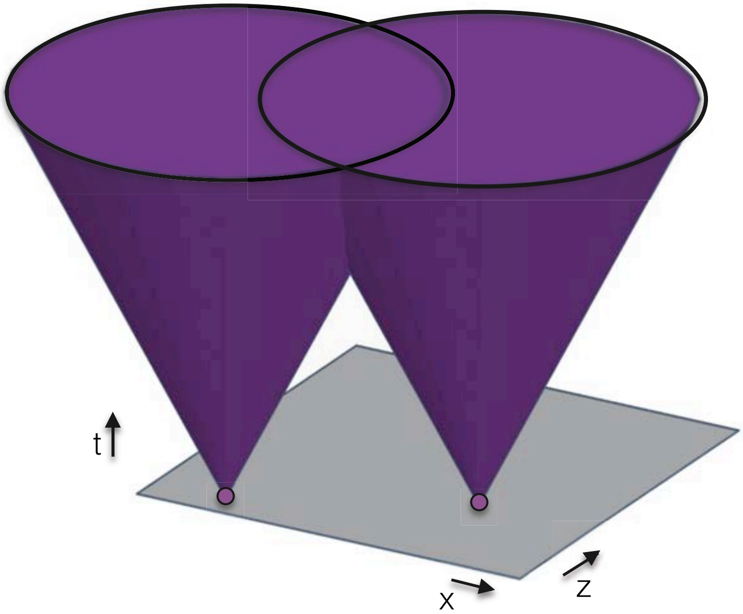


3D space-time



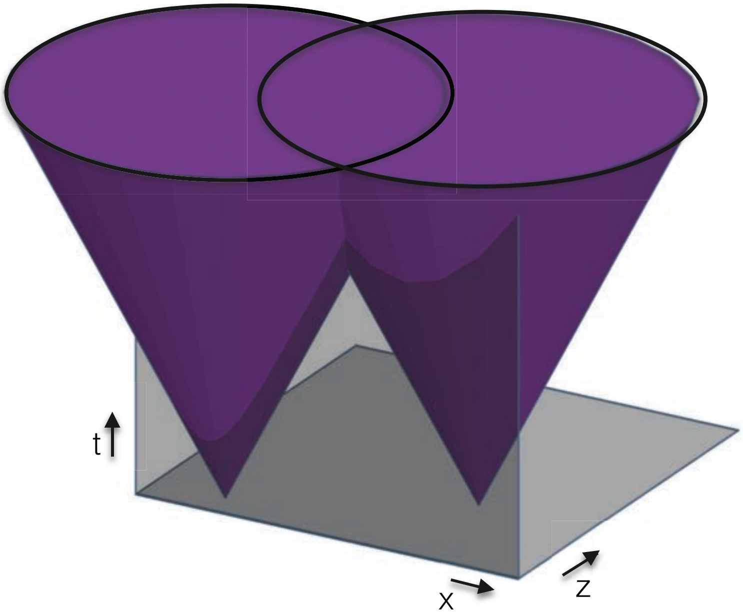
Minkowski's light cone

Transient Light Transport Analysis



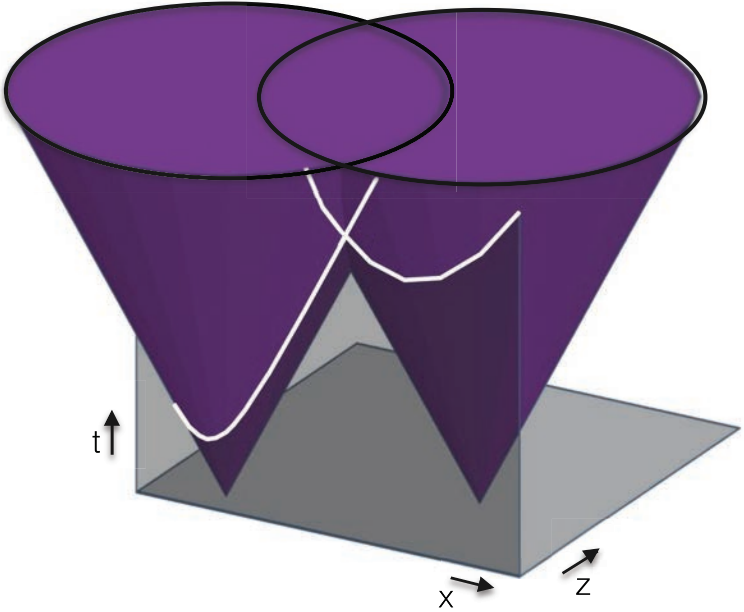
shift-invariant convolution with light cone

Transient Light Transport Analysis



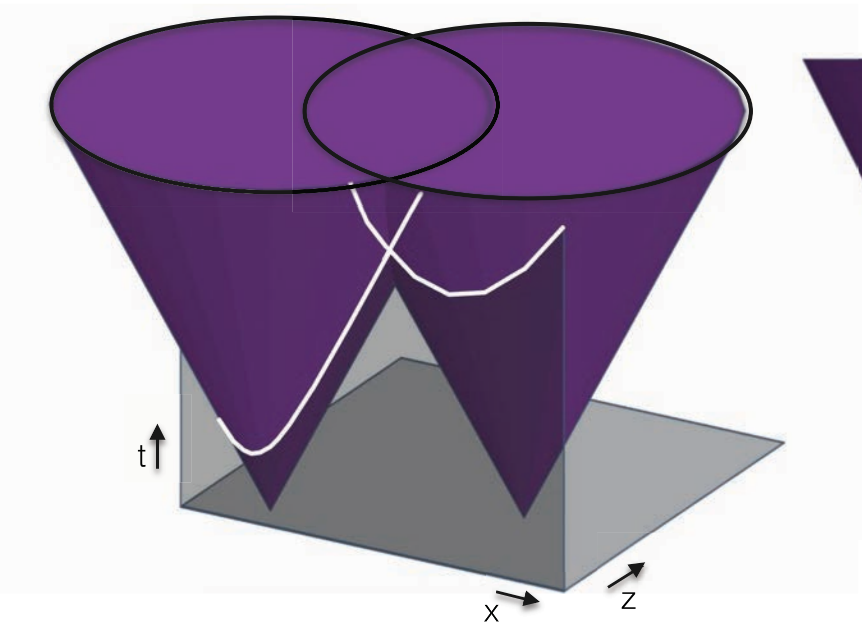
spatio-temporal information transfer

Transient Light Transport Analysis

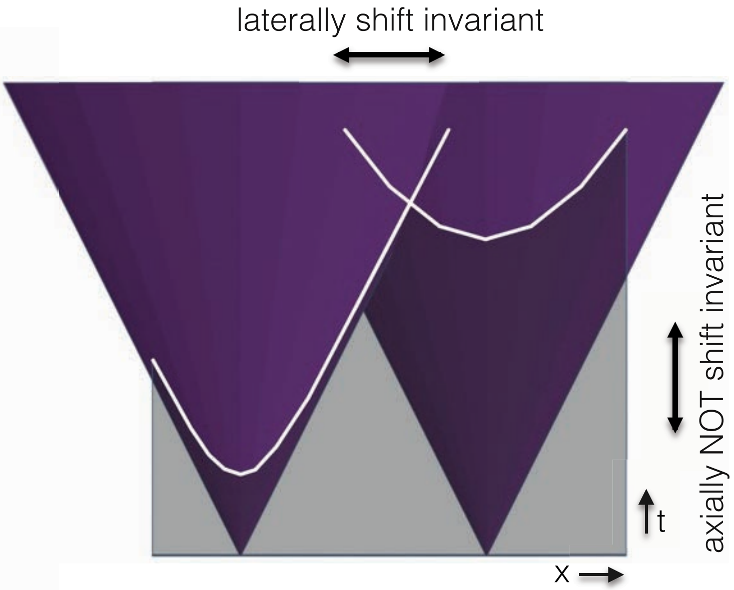


intersection with light cone

Transient Light Transport Analysis

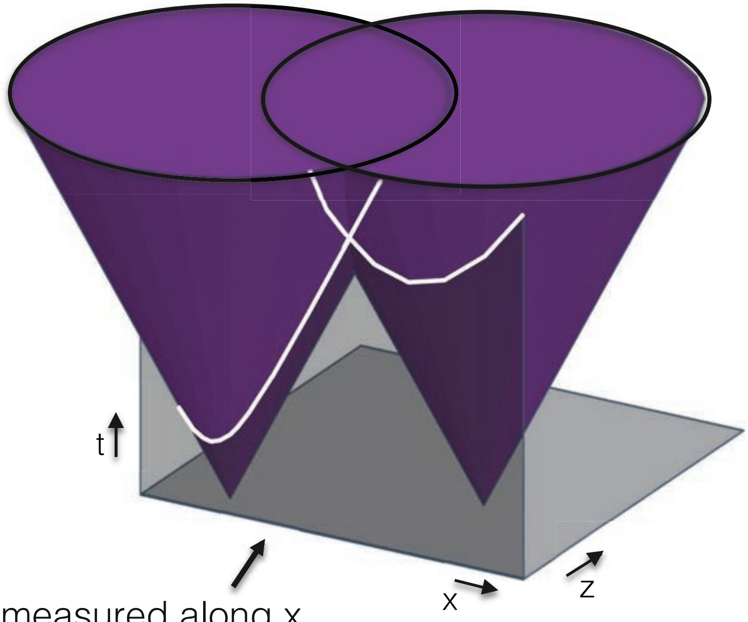


intersection with light cone

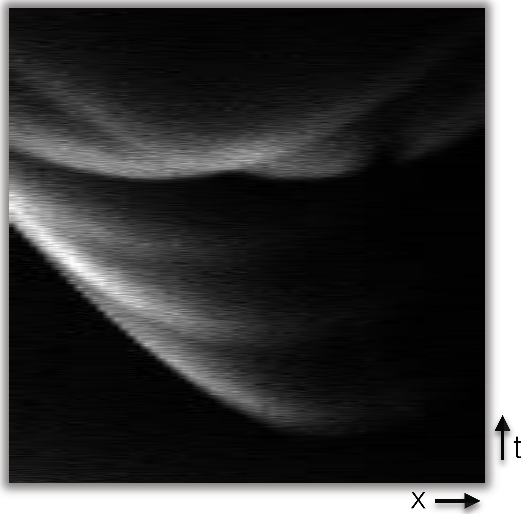


convolution – NOT shift invariant

Transient Light Transport Analysis

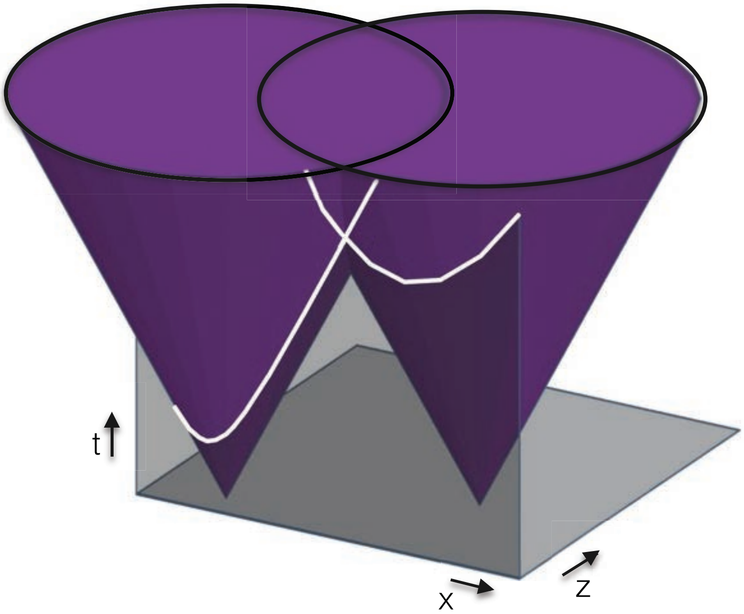


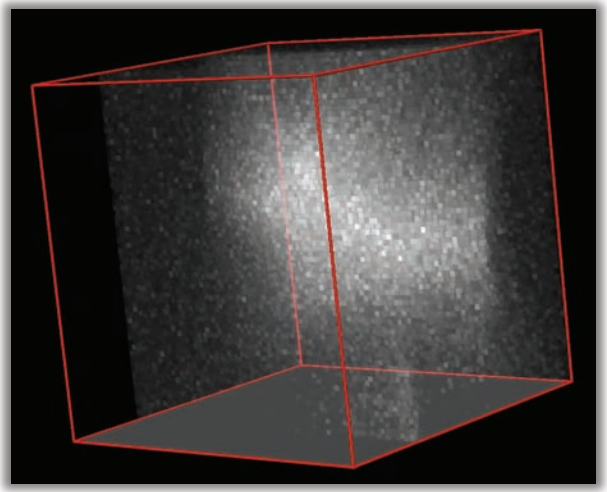
measured along x



2D SPAD measurements

Transient Light Transport Analysis



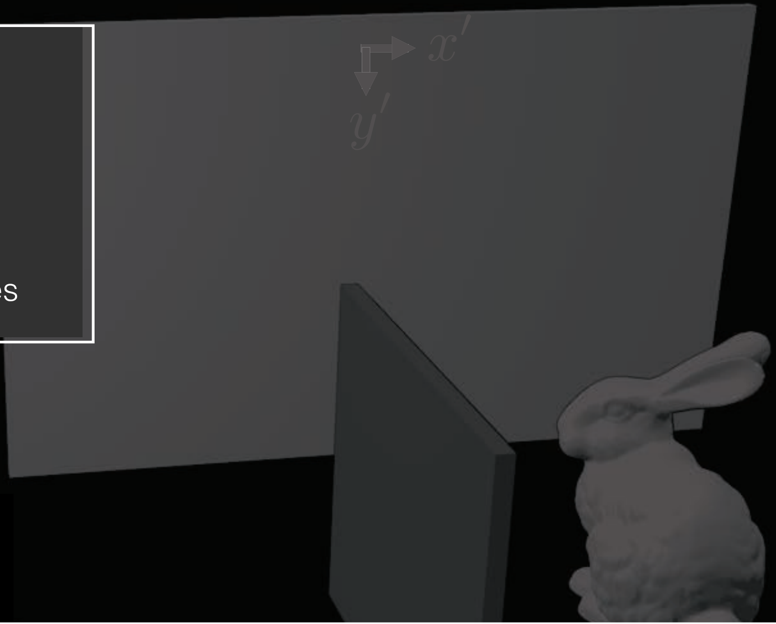


3D SPAD measurements

Our approach

express image formation model as a 3D convolution, by:

1. confocalizing measurements
2. performing a change of variables



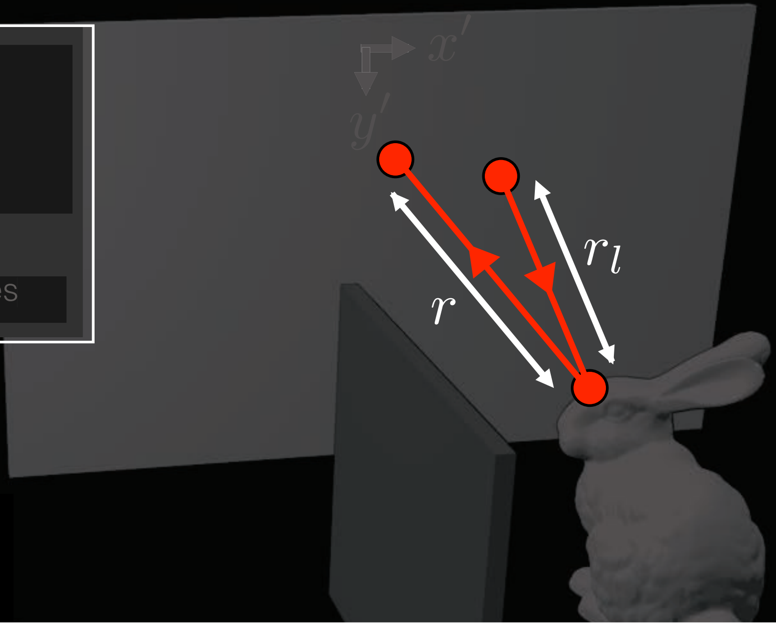
$$\tau(x', y', t) = \iiint_{\Omega} \frac{1}{r_l^2 r^2} \delta\left(r_l + r - tc\right) \cdot \rho(x, y, z)\ dx\ dy\ dz$$

Our approach

express image formation model as a 3D convolution, by:

1. confocalizing measurements
2. performing a change of variables

3D measurements

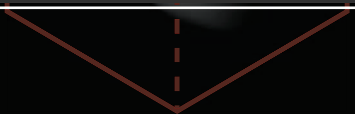


$$\tau(x',y',t) = \iiint_{\Omega} \frac{1}{r_l^2 r^2} \delta\left(r_l + r - tc\right) \cdot \rho(x,y,z)\ dx\ dy\ dz$$

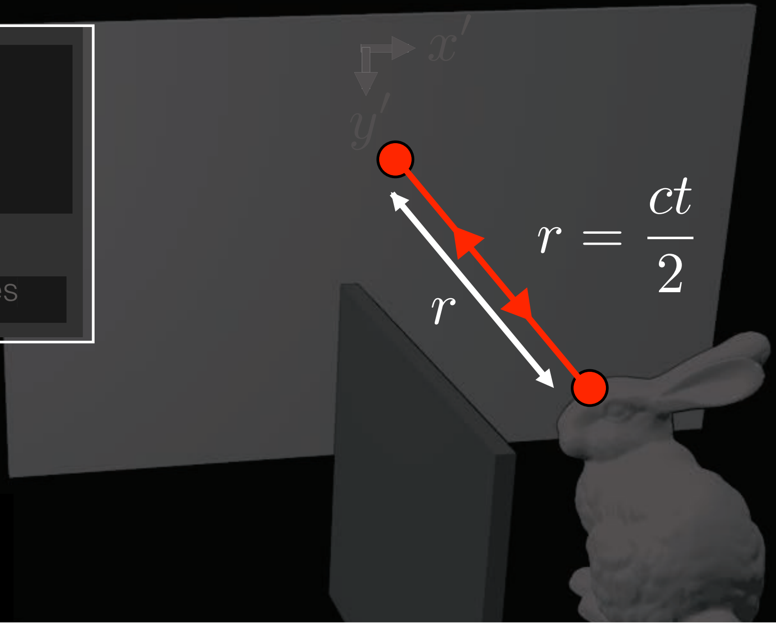
Our approach

express image formation model as a 3D convolution, by:

1. confocalizing measurements
2. performing a change of variables



3D measurements



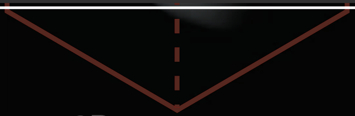
$$\tau(x', y', t) = \iiint_\Omega \frac{1}{r^4} \delta\left(2r - tc\right) \cdot \rho(x, y, z)\ dx\ dy\ dz$$

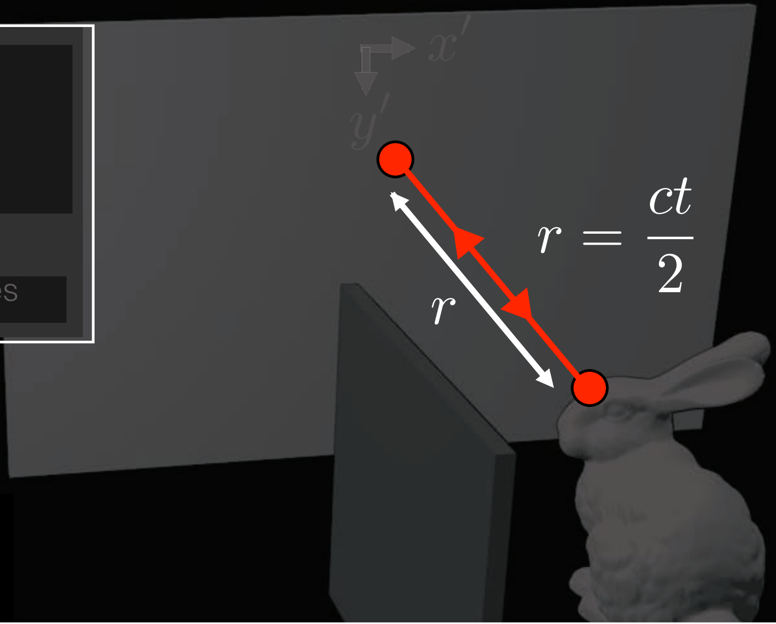
hypercone with factor 2

Our approach

express image formation model as a 3D convolution, by:

1. confocalizing measurements
2. performing a change of variables





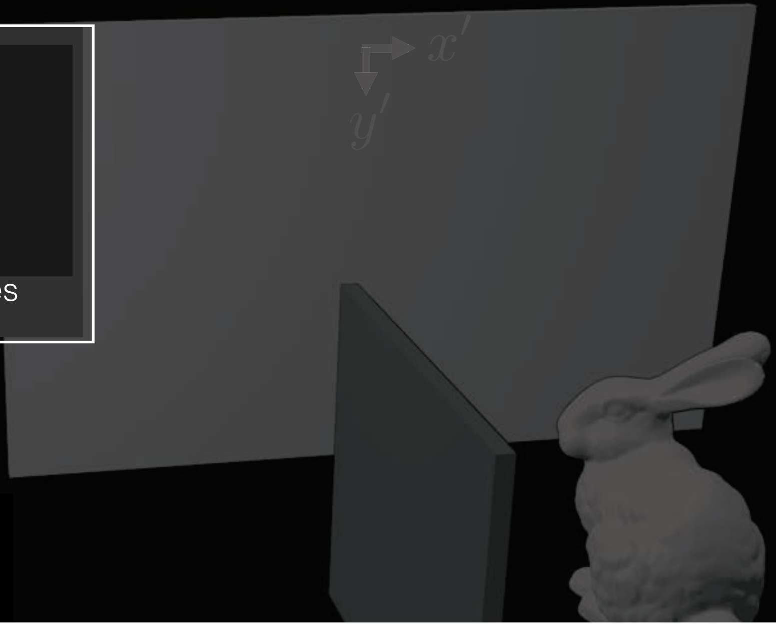
$$\tau(x', y', t) = \iiint_{\Omega} \frac{1}{r^4} \delta(2r - tc) \cdot \rho(x, y, z)\ dx\ dy\ dz$$

$$\frac{1}{r^4} = \left(\frac{2}{ct}\right)^4$$

Our approach

express image formation model as a 3D convolution, by:

1. confocalizing measurements
2. performing a change of variables

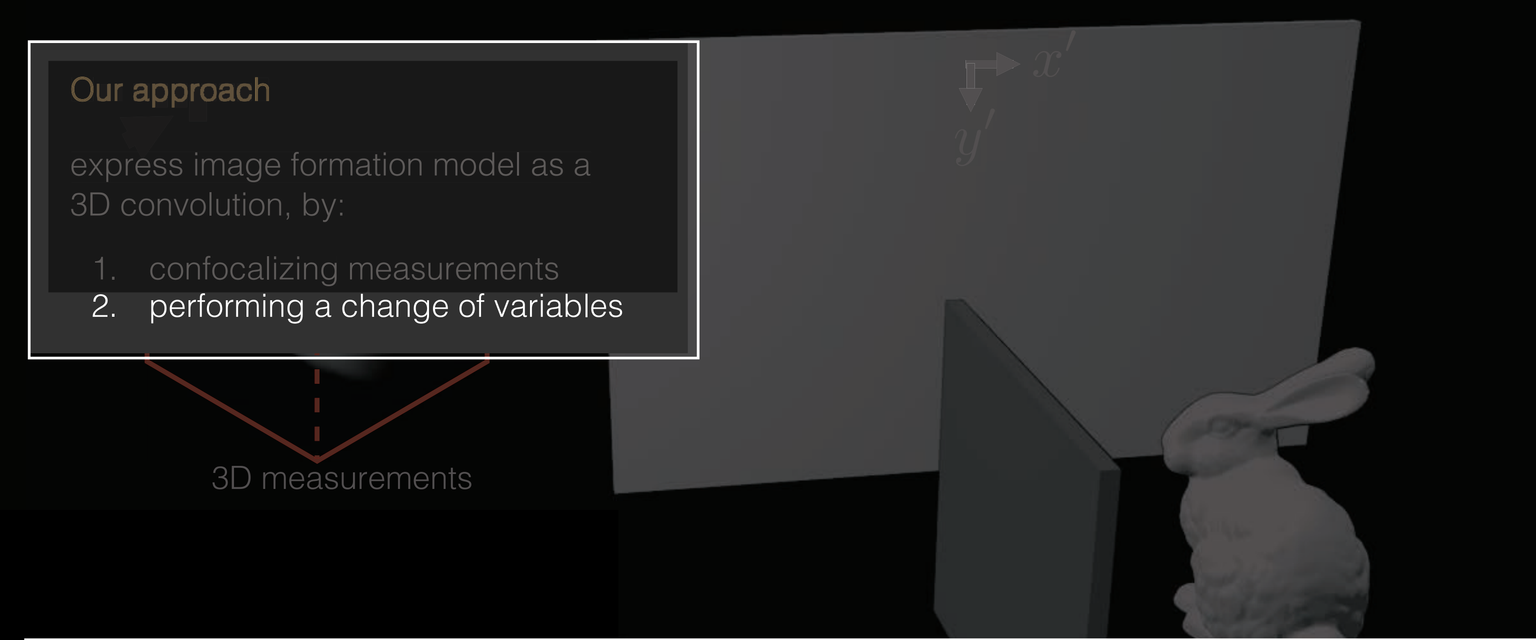


$$\tau(x', y', t) = \left(\frac{2}{ct}\right)^4 \iiint_\Omega \delta\left(2r - tc\right) \cdot \rho(x, y, z)\; dx\; dy\; dz$$

Our approach

express image formation model as a 3D convolution, by:

1. confocalizing measurements
2. performing a change of variables



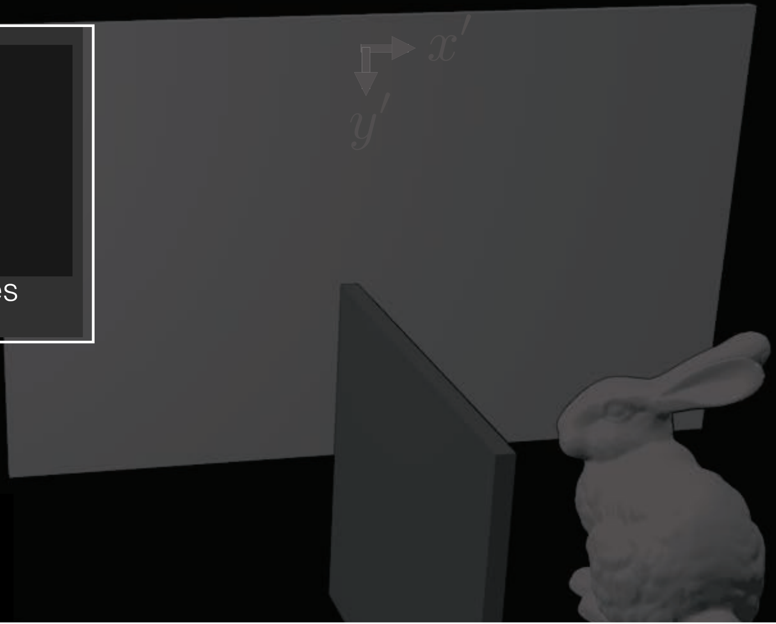
$$v^{3/2}\tau(x', y', \frac{2}{c}\sqrt{v}) = \iiint_\Omega \delta\left((x'-x)^2 + (y'-y)^2 + u - v\right) \cdot \frac{1}{2\sqrt{u}}\, \rho(x, y, \sqrt{u})\, dx\, dy\, du$$

Our approach

express image formation model as a 3D convolution, by:

1. confocalizing measurements
2. performing a change of variables

3D measurements



$$v^{3/2}\tau(x', y', \frac{2}{c}\sqrt{v}) = \iiint_{\Omega} \delta\left((x'-x)^2 + (y'-y)^2 + u - v\right) \cdot \frac{1}{2\sqrt{u}}\, \rho(x, y, \sqrt{u})\, dx\, dy\, du$$

$$\tau = \mathbf{a} * \rho$$

NLOS image formation mode:

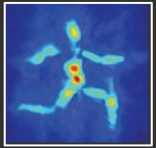
$$\tau = \mathbf{A}\rho$$

measurements	transport matrix	unknown volume
$n^3 \times 1$	$n^3 \times n^3$	$n^3 \times 1$

Backpropagation [Velten 12, Buttafava 15]

Flops: $O(n^5)$

Memory: $O(n^3)$



Iterative Inversion [Gupta 12, Wu 12, Heide 13]

Flops: $O(n^5)$ per iter.

Memory: $O(n^5)$

NLOS image formation mode:

$$\tau = \mathbf{A}\rho$$

measurements $n^3 \times 1$

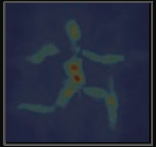
transport matrix $n^3 \times n^3$

unknown volume $n^3 \times 1$

Backpropagation [Velten 12, Buttafava 15]

Flops: $O(n^5)$

Memory: $O(n^3)$



Iterative Inversion [Gupta 12, Wu 12, Heide 13]

Flops: $O(n^5)$ per iter.

Memory: $O(n^5)$

Confocal NLOS image formation mode:

$$\tau = \mathbf{a} * \rho$$

measurements $n \times n \times n$

blur kernel $n \times n \times n$

unknown volume $n \times n \times n$

NLOS image formation mode:

$$\tau = \mathbf{A}\rho$$

measurements $n^3 \times 1$

transport matrix $n^3 \times n^3$

unknown volume $n^3 \times 1$

Confocal NLOS image formation mode:

$$\tau = \mathbf{a} * \rho$$

measurements $n \times n \times n$

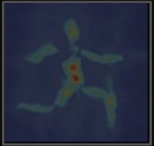
blur kernel $n \times n \times n$

unknown volume $n \times n \times n$

Backpropagation [Velten 12, Buttafava 15]

Flops: $O(n^5)$

Memory: $O(n^3)$



Iterative Inversion [Gupta 12, Wu 12, Heide 13]

Flops: $O(n^5)$ per iter.

Memory: $O(n^5)$

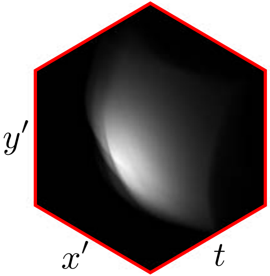
3D Deconvolution with Light Cone Transform

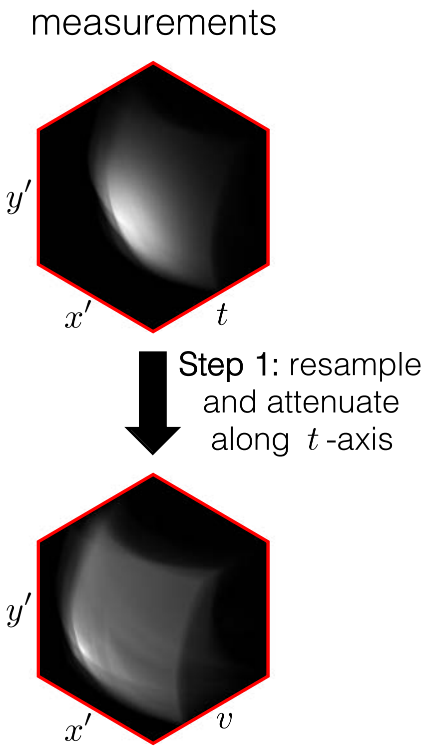
[O'Toole et al. 2018, Nature (to appear)]

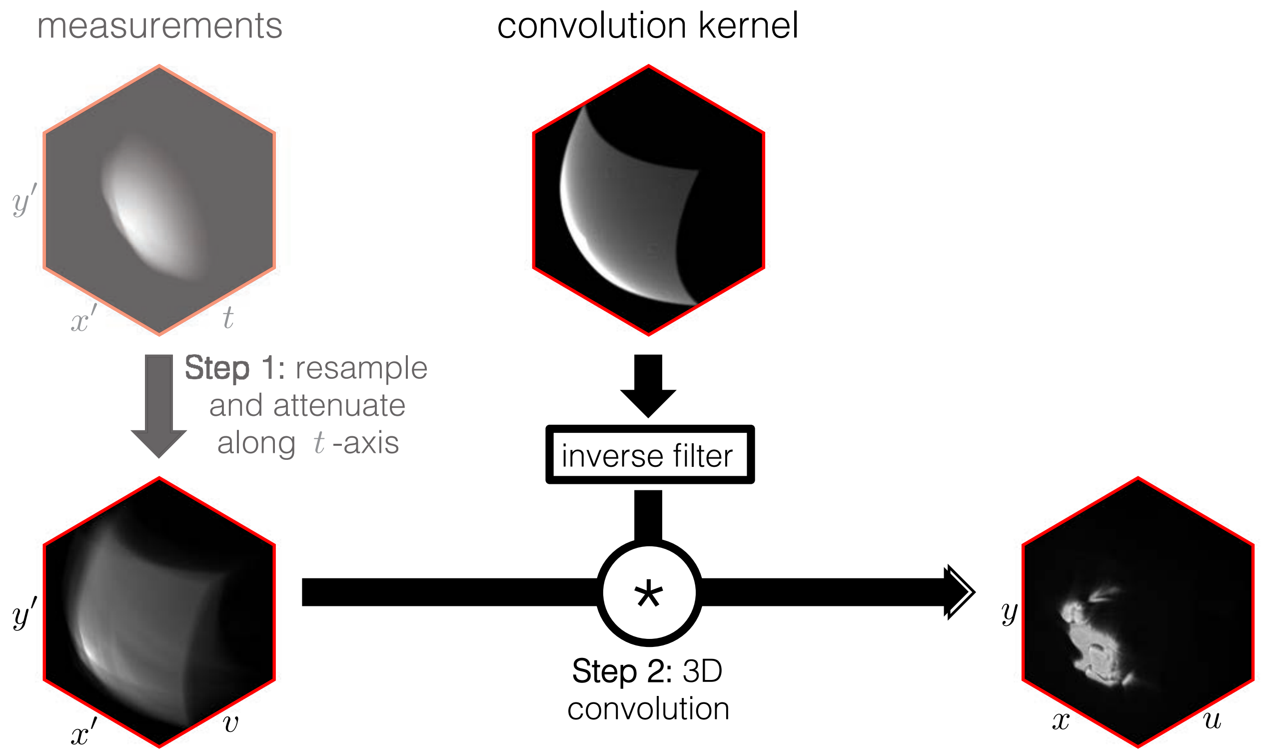
Flops: $O(n^3 \log(n))$

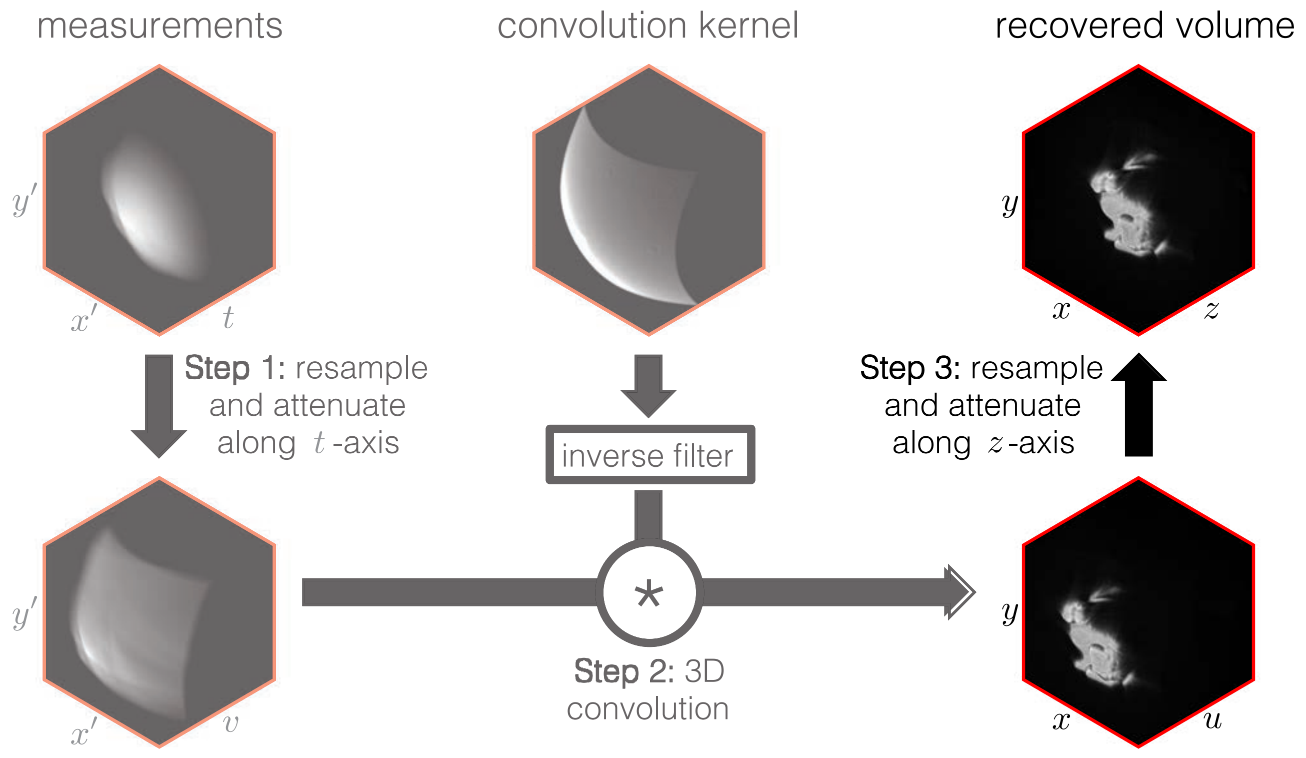
Memory: $O(n^3)$

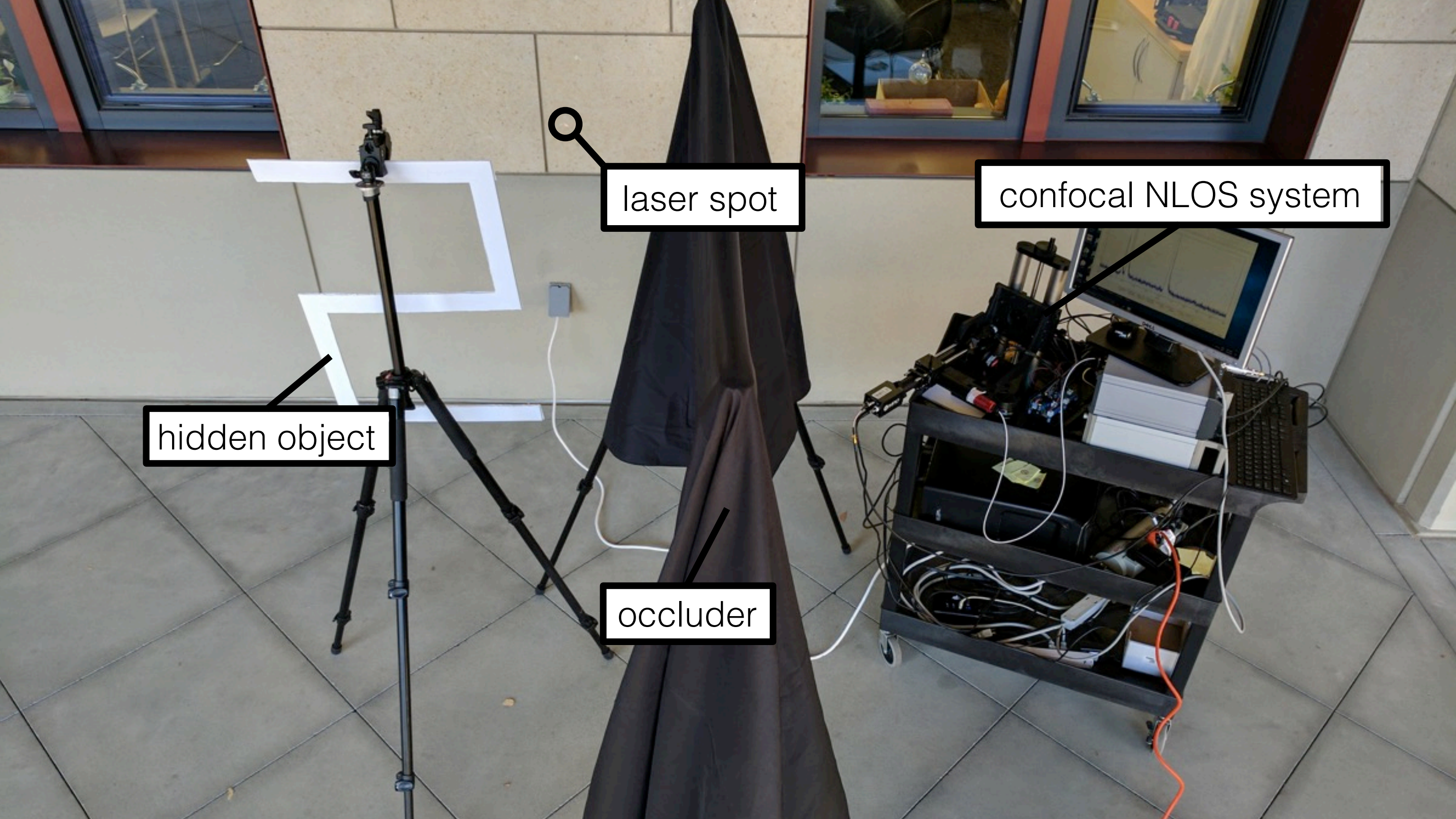
measurements




measurements
y′
x′
t
Step 1: resample
and attenuate
along t -axis
y′
x′
v

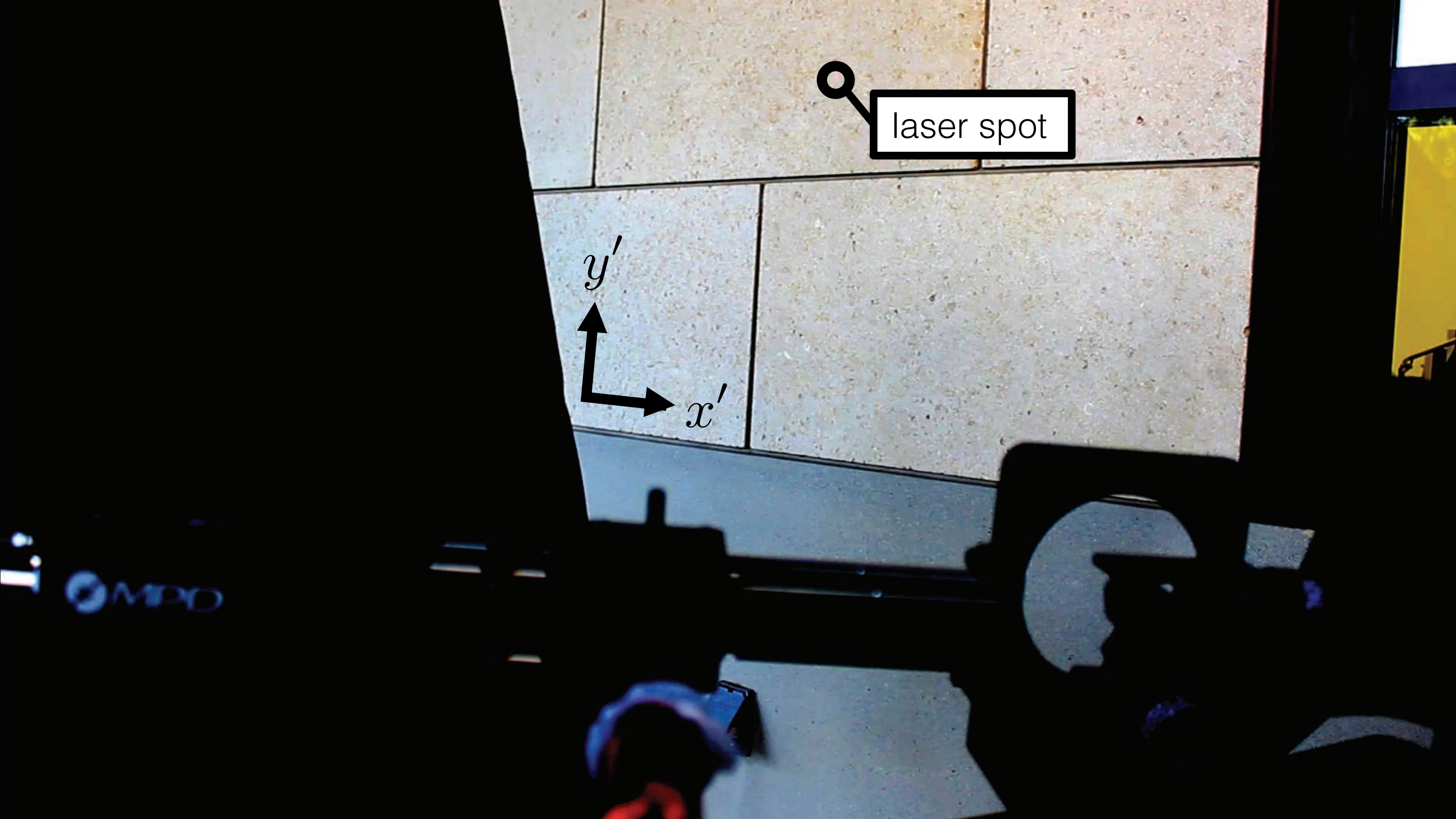

measurements
convolution kernel
y'
x'
t
Step 1: resample and attenuate along t -axis
inverse filter
y'
x'
v
*
Step 2: 3D convolution
y
x
u


measurements
convolution kernel
recovered volume
y'
x'
t
y
x
z
Step 1: resample and attenuate along t-axis
inverse filter
Step 3: resample and attenuate along z-axis
y'
x'
v
*
Step 2: 3D convolution
y
x
u

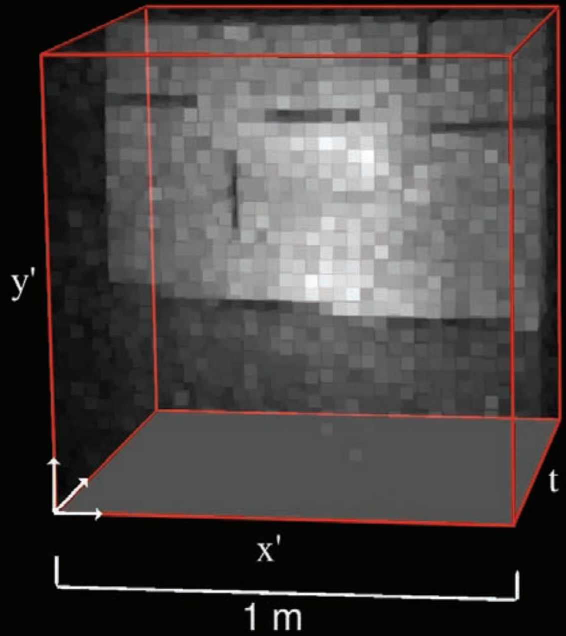


laser spot
confocal NLOS system
hidden object
occluder


1.25 meters


laser spot
y'
x'

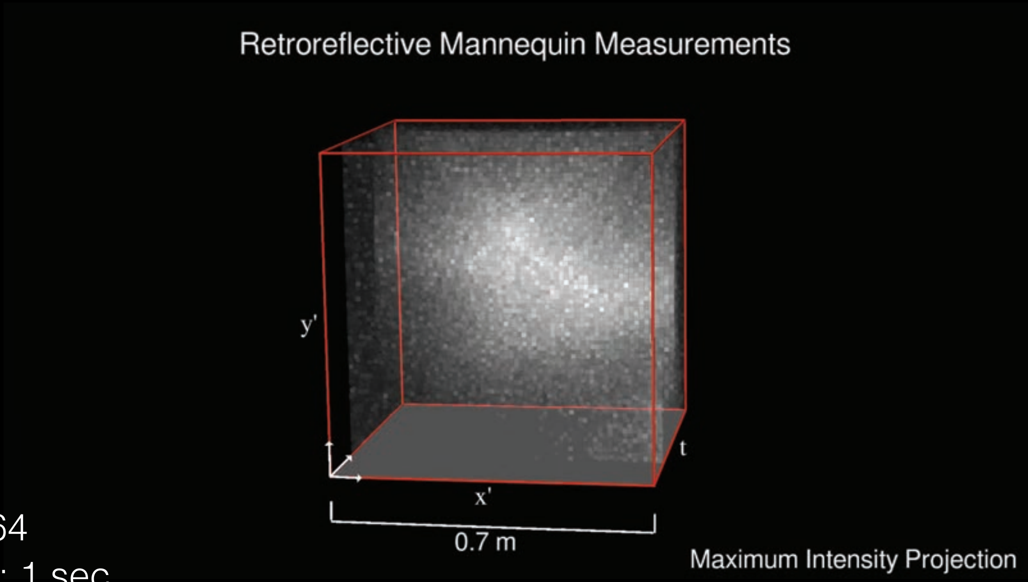
measurements



Maximum Intensity Projection

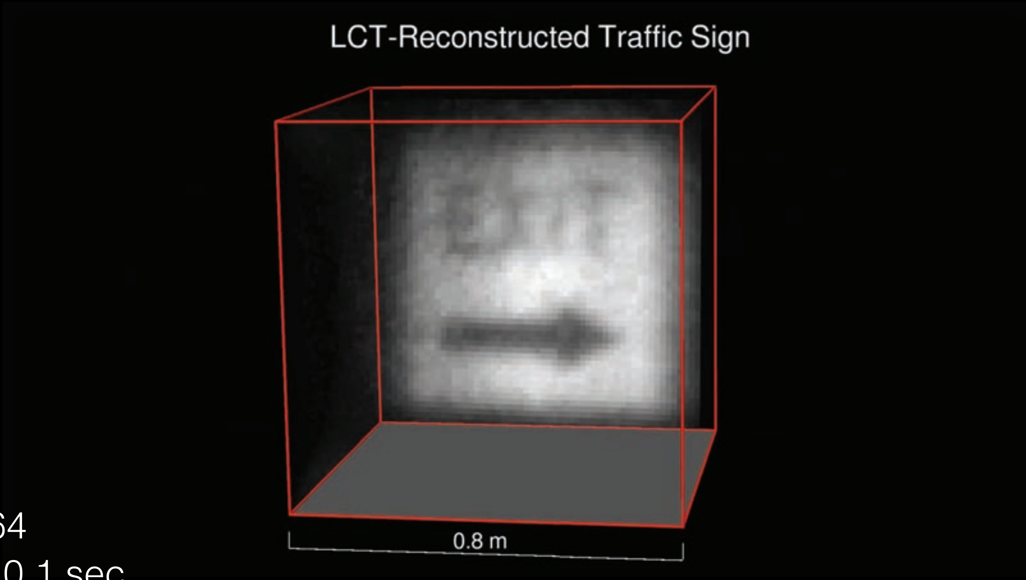
Experiments

Spatial resolution: 64x64
Exposure time (per sample): 1 sec
Retroreflective: Yes

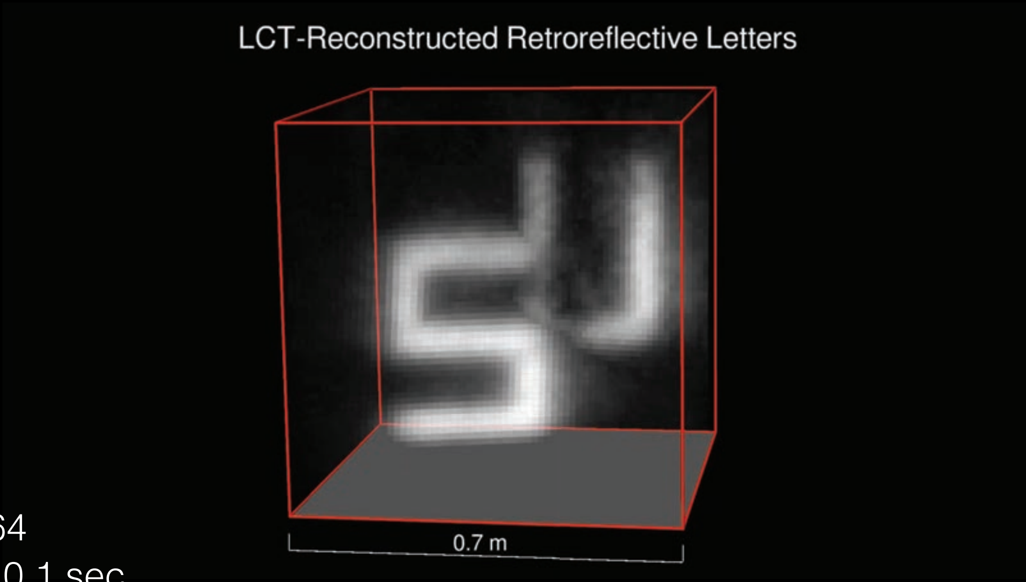


Experiments

Spatial resolution: 64x64
Exposure time (per sample): 0.1 sec
Retroreflective: Yes

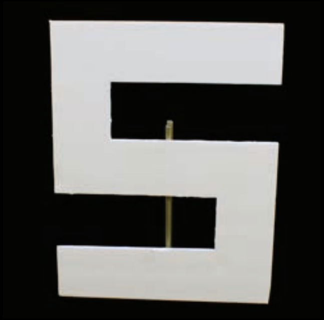


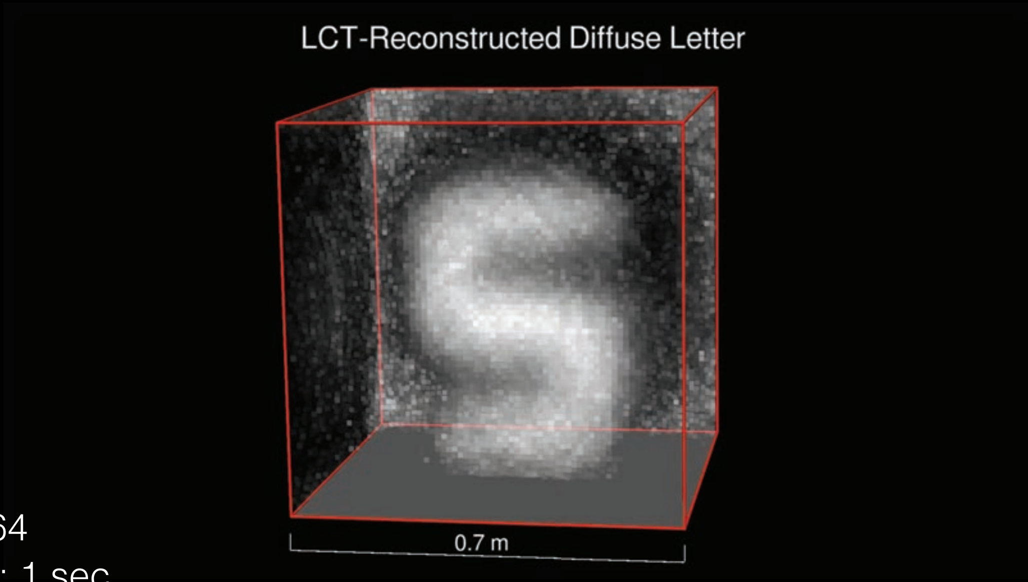
Experiments



Spatial resolution: 64x64
Exposure time (per sample): 0.1 sec
Retroreflective: Yes

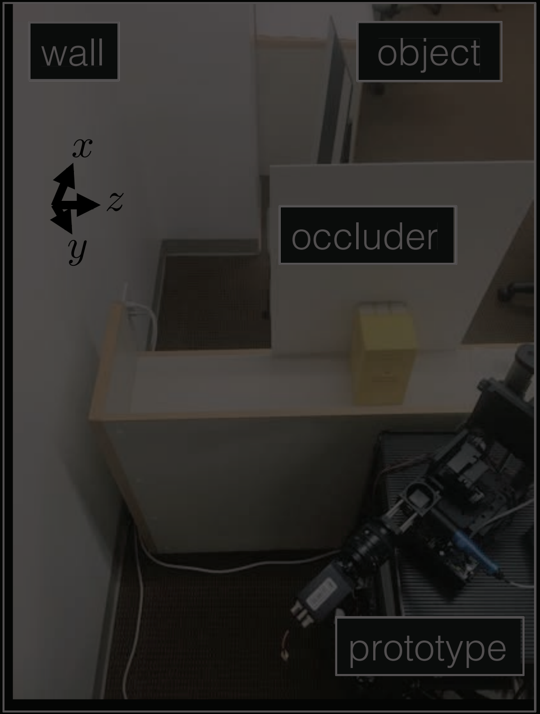
Experiments

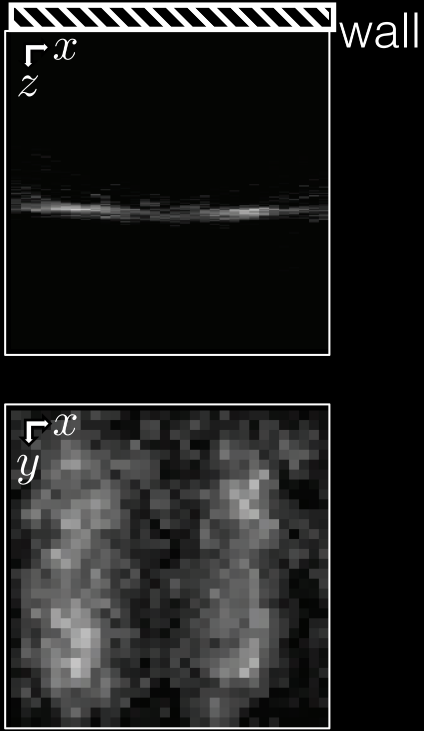




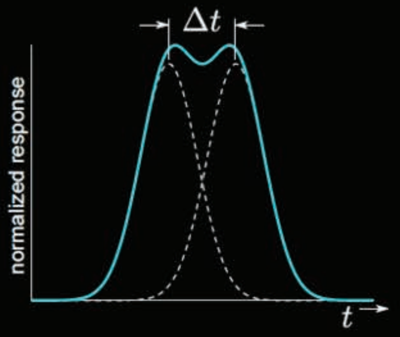
Spatial resolution: 64x64
Exposure time (per sample): 1 sec
Retroreflective: No

Towards Real-time NLOS Imaging





Resolution Limits of NLOS Imaging



$$\Delta x \geq \frac{c \cdot \sqrt{w^2 + z^2}}{2w} \text{FWHM}$$

Challenges of NLOS Imaging

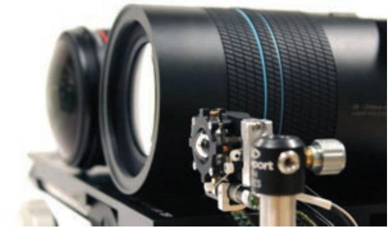
1. Light efficiency, high-speed time stamping
2. Efficient Scanning
3. Large-scale inverse problem
4. Accurate (and invertible) NLOS light transport model

… slides omitted for confidentiality …

Stanford Computational Imaging Lab

Virtual & Augmented Reality

Light Field Cameras



Computational Microscopy

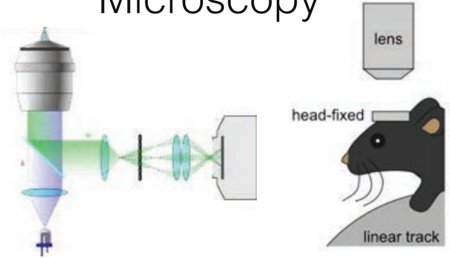
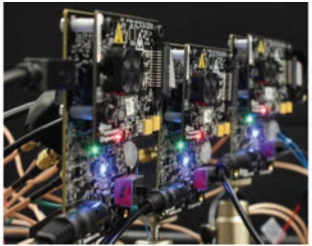


Image Optimization

Time-of-Flight Imaging



Computational Displays

Gordon Wetzstein
Computational Imaging Lab
Stanford University

www.computationalimaging.org

Matt O'Toole

David Lindell

Felix Heide

Thanks to:

- Edoardo Charbon, EPFL
- Samuel Burri, EPFL
- Pierre-Yves Cattin, Fastree3D