Non-line-of-sight imaging

15-463, 15-663, 15-862
Computational Photography
Fall 2018, Lecture 23

http://graphics.cs.cmu.edu/courses/15-463

Course announcements

- Homework 6 posted, due November 30th.
  - Three-week homework.
  - Do not leave for last minute, you won't have time to complete it.
  - Homework 7 will be appropriately shorter so that it can be done in the last week.

- Doodle for scheduling final project presentations was posted on Piazza.
  - https://doodle.com/poll/zf4ur49692m772eg
  - Please make sure to vote!

- Project checkpoint meetings next week.
  - I will post a spreadsheet so that you can sign up for a meeting slot.

Overview of today’s lecture

- The non-line-of-sight (NLOS) imaging problem.
- Active NLOS imaging using time-of-flight imaging.
- Active NLOS imaging using WiFi.
- Passive NLOS imaging using accidental pinholes.
- Passive NLOS imaging using accidental reflectors.
- Passive NLOS imaging using corners.

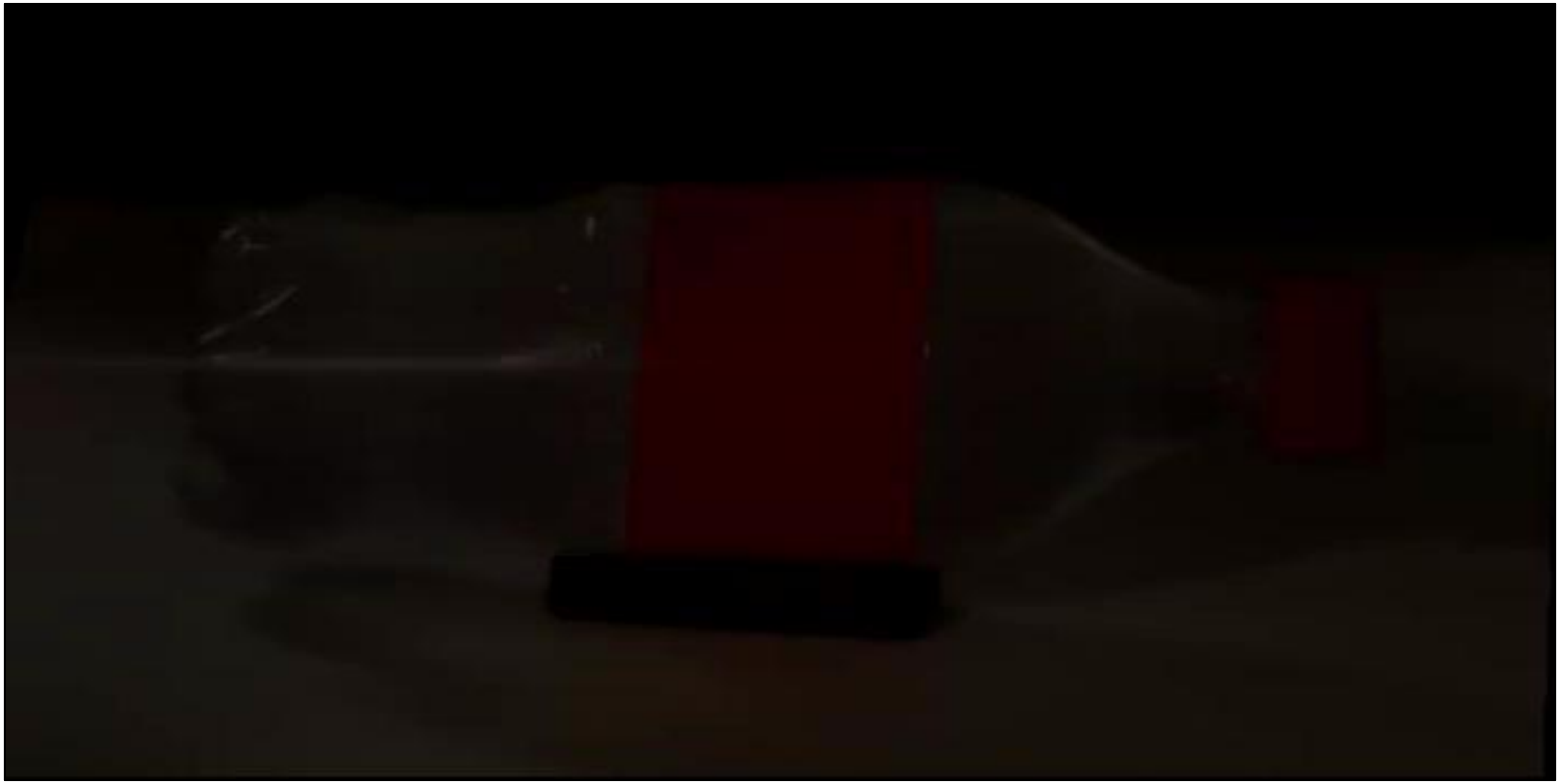
Slide credits

Many of these slides were directly adapted from:

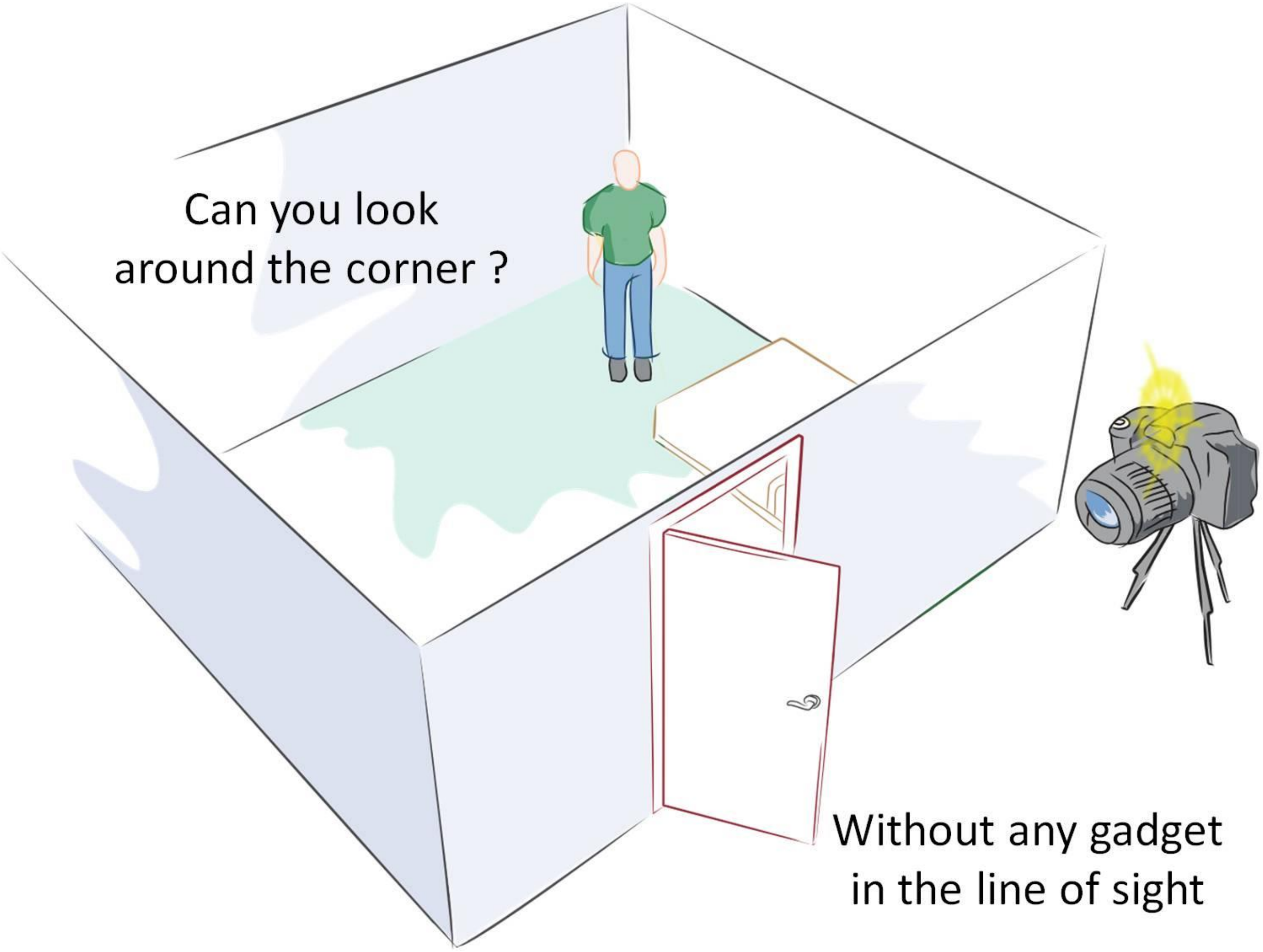
- Matthew O’Toole (CMU).
- Shree Nayar (Columbia).
- Fadel Adib (MIT).
- Katie Bouman (MIT).
- Adithya Pediredla (Rice).

The non-line-of-sight (NLOS) imaging problem

Time-of-flight (ToF) imaging



Looking around the corner



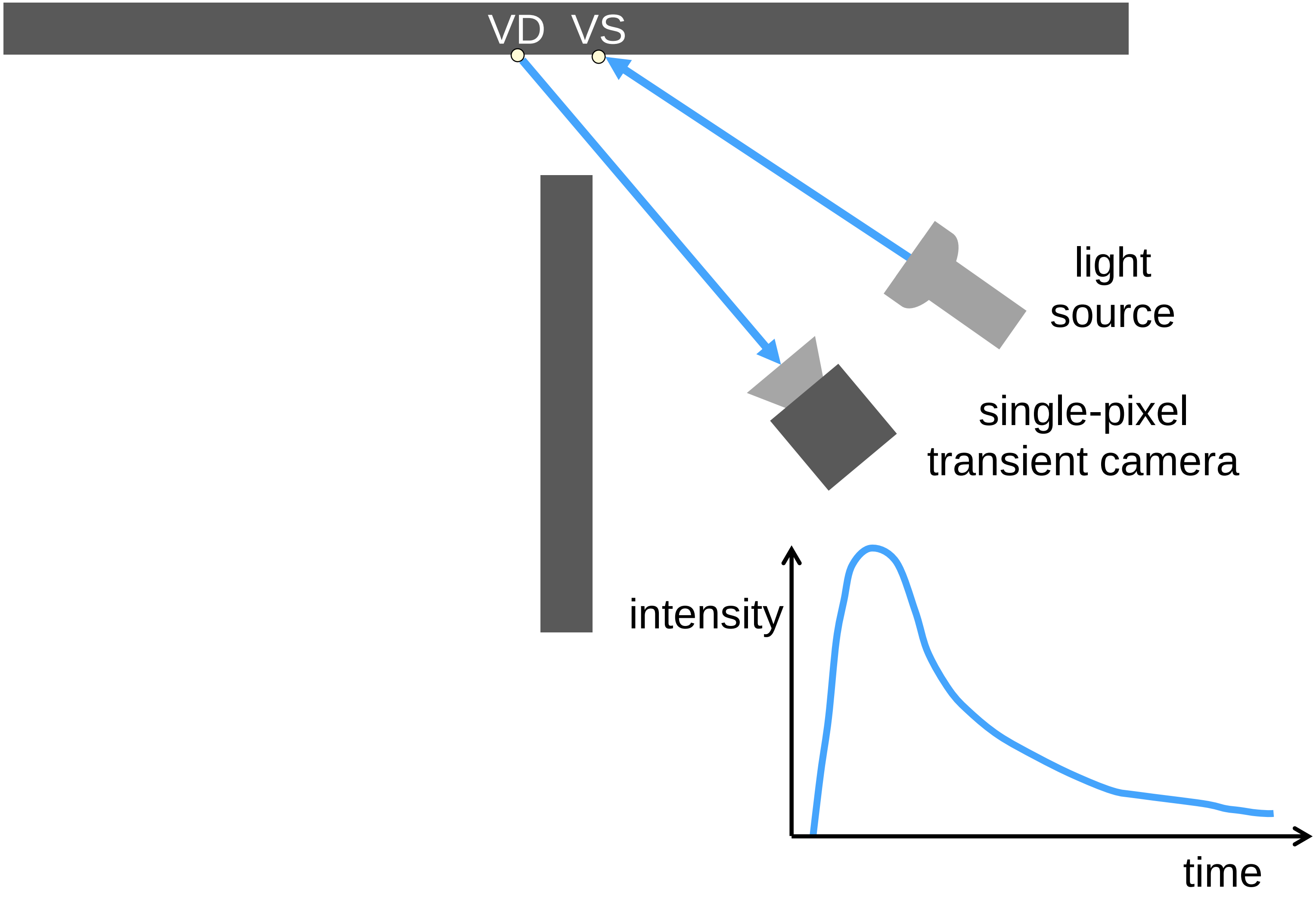
Can you look around the corner ?

Without any gadget in the line of sight

Looking around the corner

Active NLOS imaging using time-of-flight imaging

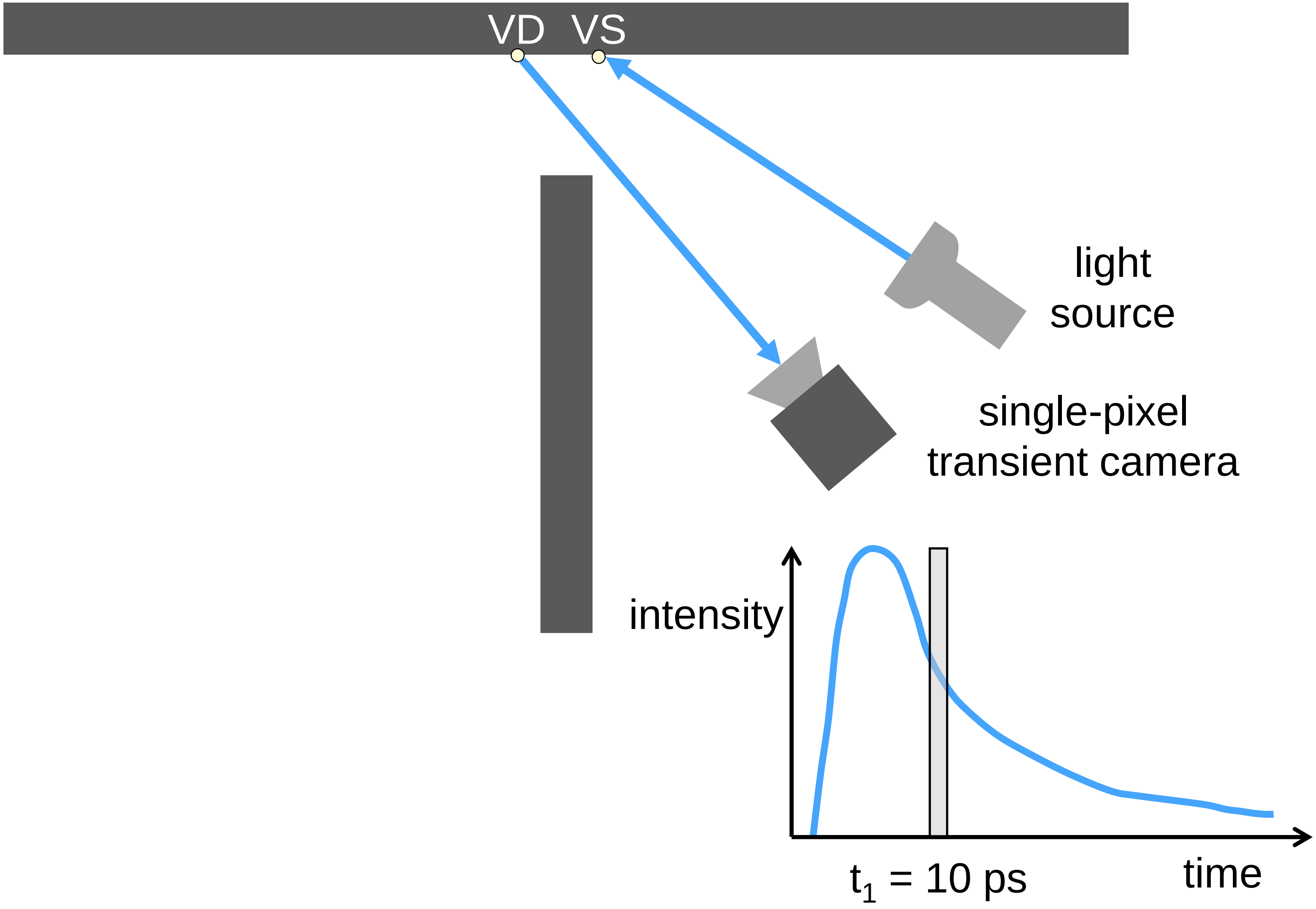
Processing a single-pixel transient



light
source

single-pixel
transient camera

Processing a single-pixel transient



VD VS

light source

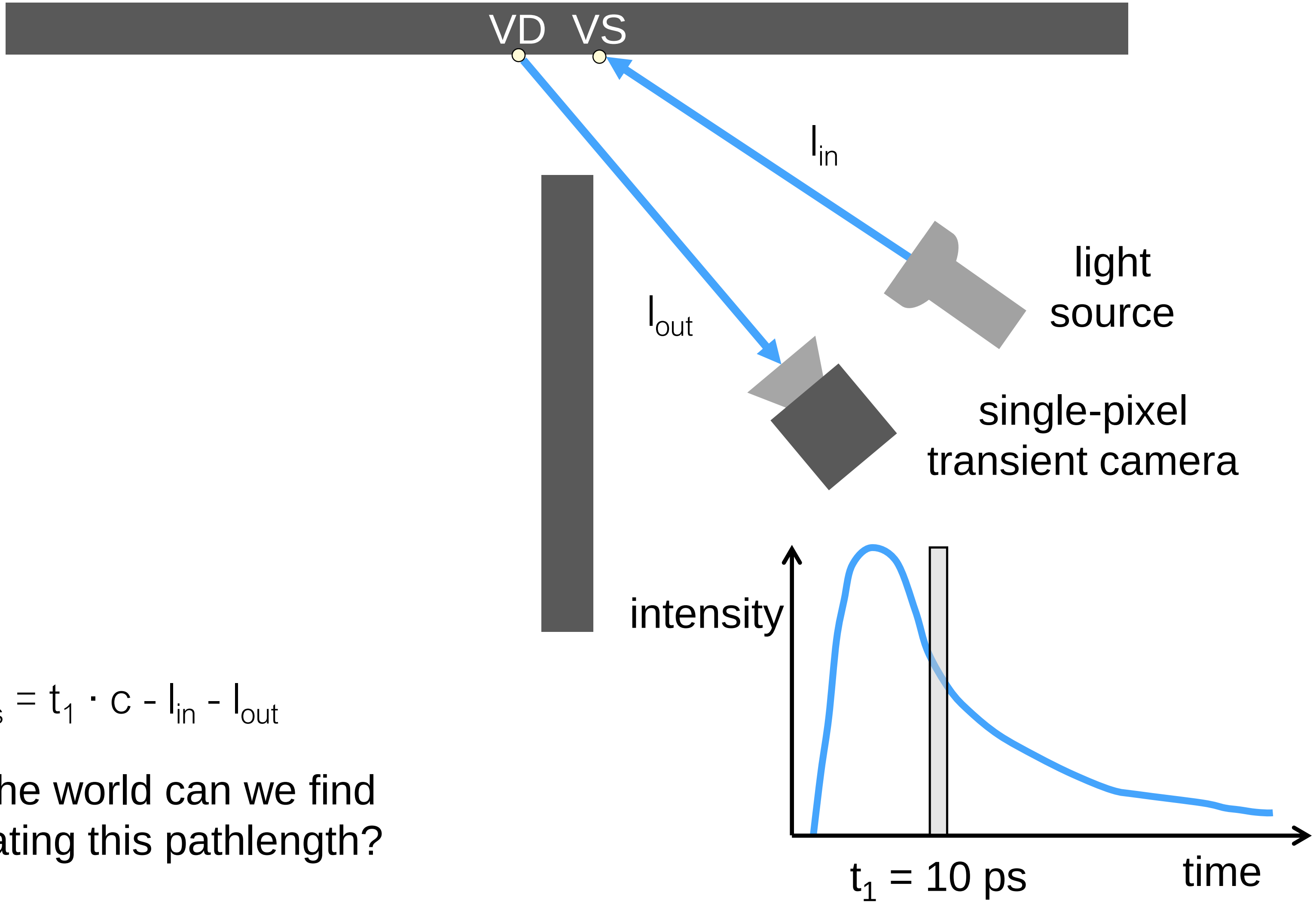
single-pixel transient camera

intensity

t_1 = 10 ps

time

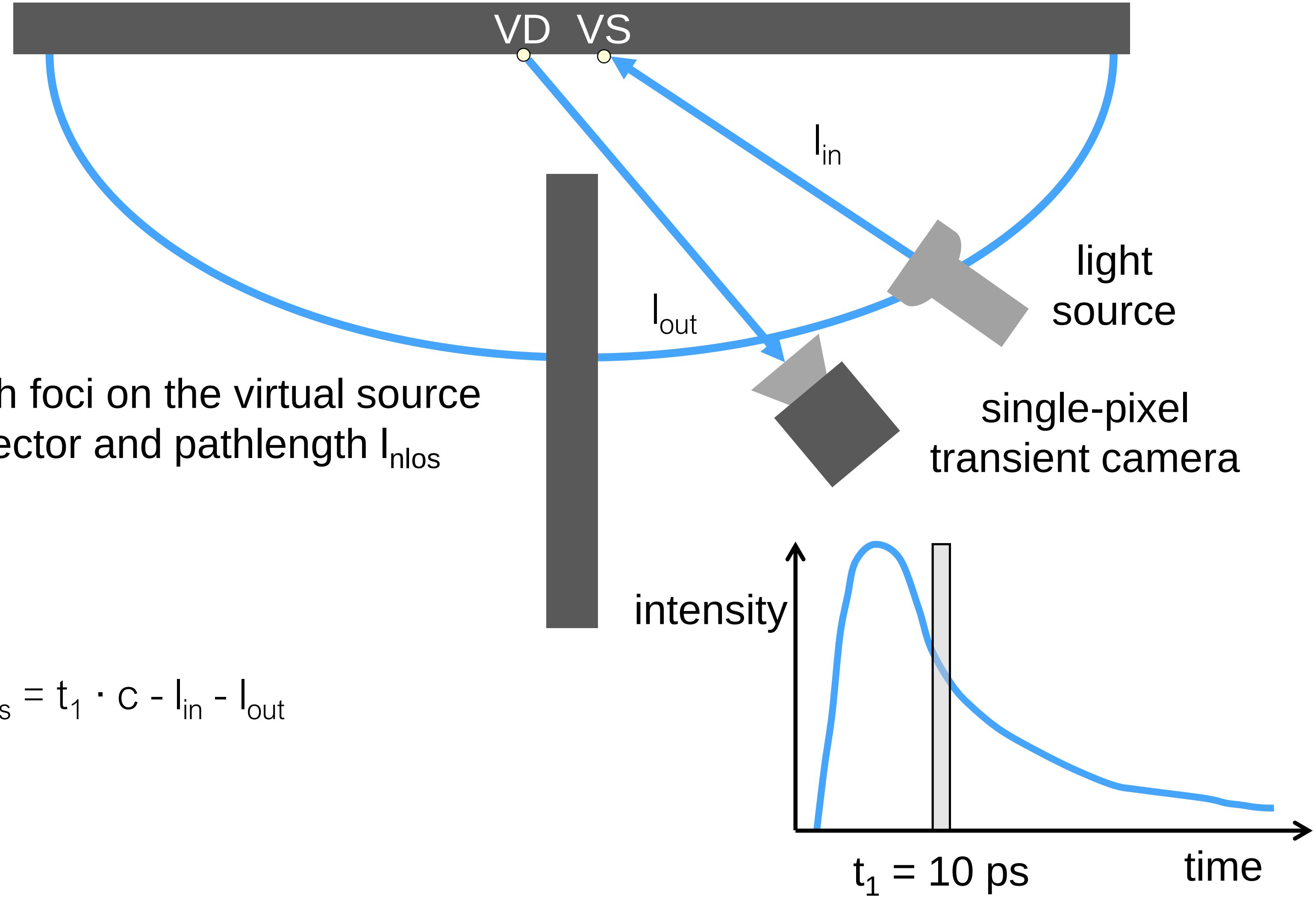
Processing a single-pixel transient



$l_{nlos} = t_1 \cdot c - l_{in} - l_{out}$

Where in the world can we find points creating this pathlength?

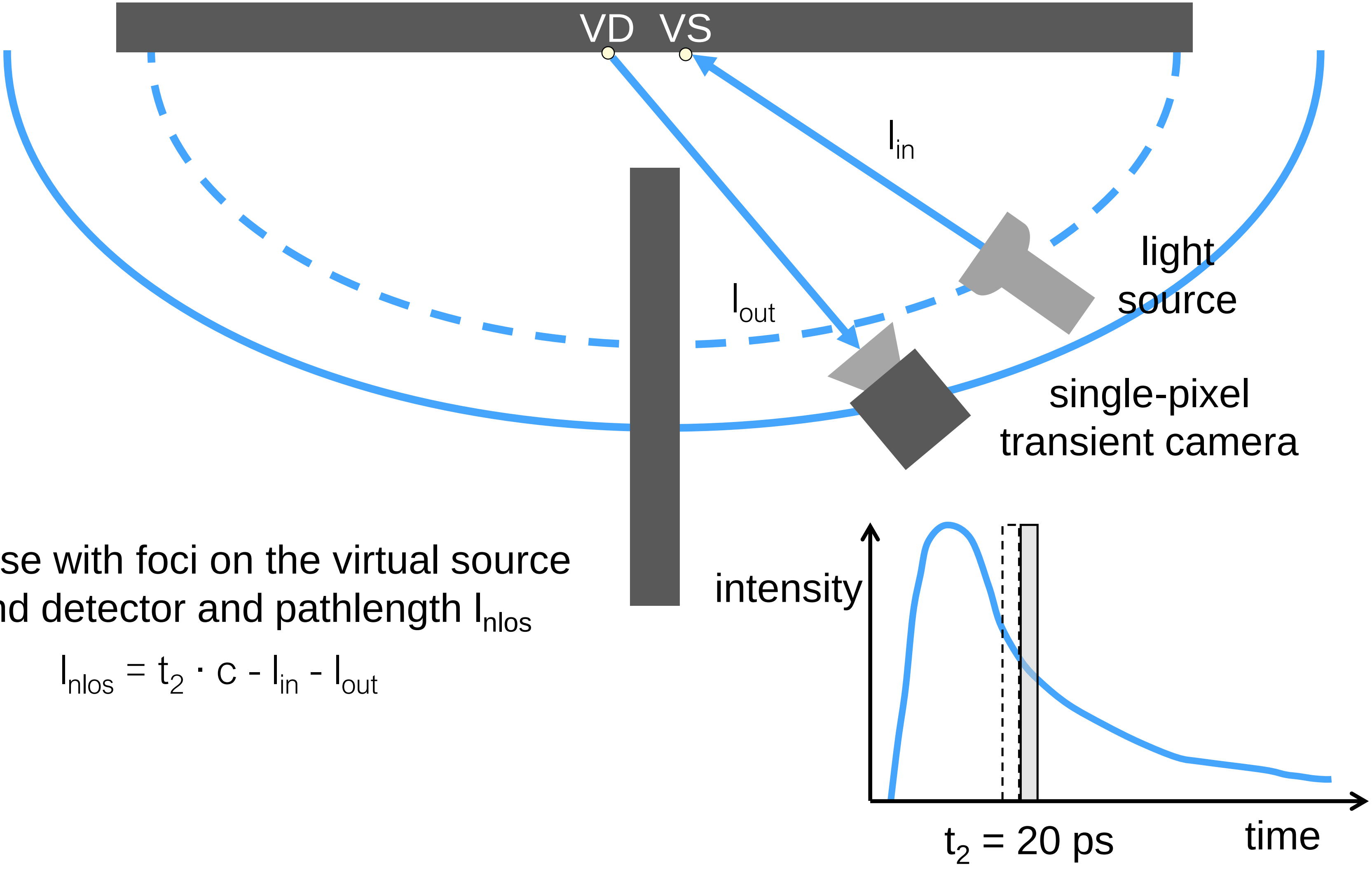
Processing a single-pixel transient



Ellipse with foci on the virtual source and detector and pathlength l_{nlos}

$$l_{nlos} = t_1 \cdot c - l_{in} - l_{out}$$

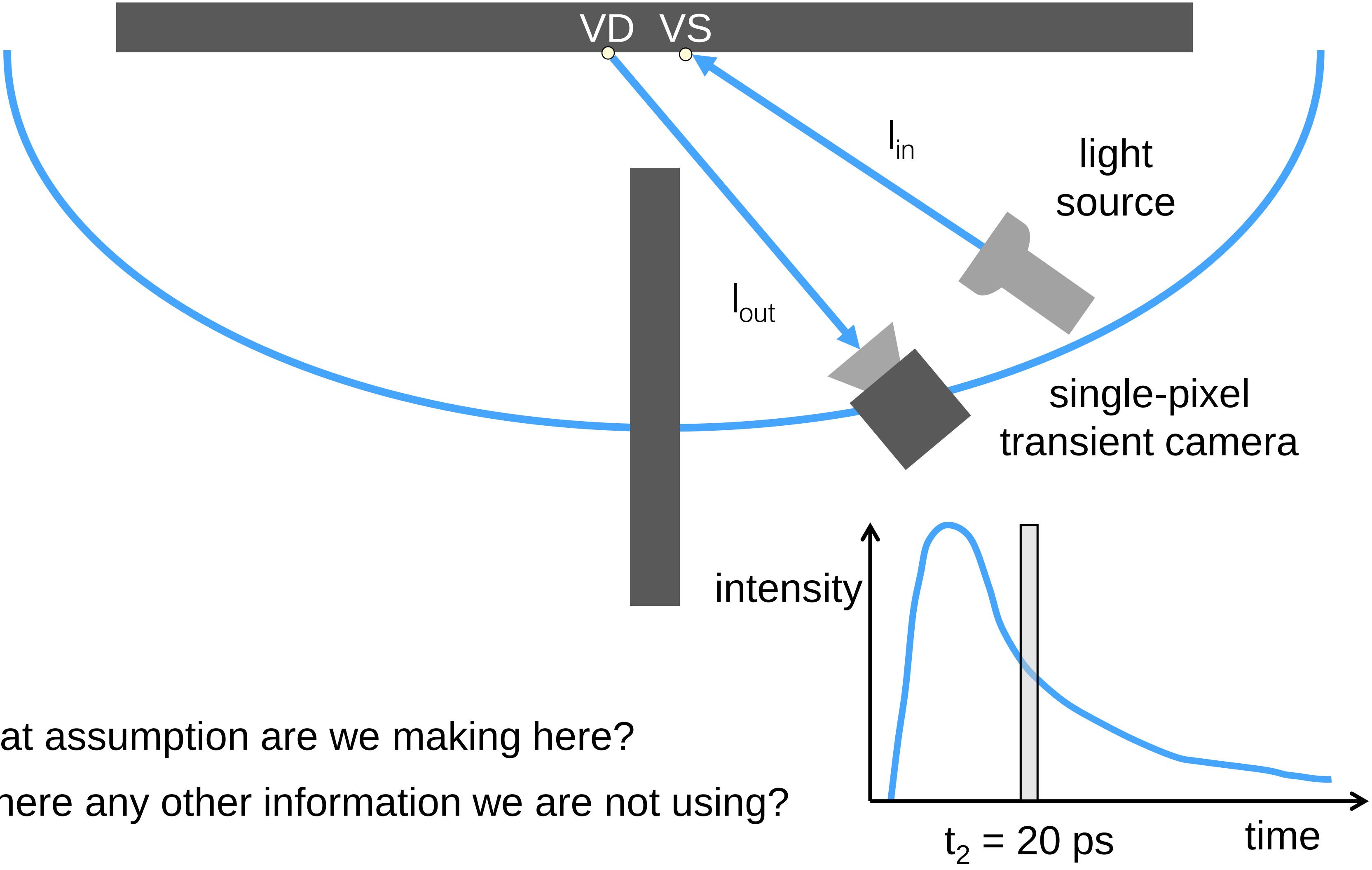
Processing a single-pixel transient



Ellipse with foci on the virtual source and detector and pathlength l_{nlos}

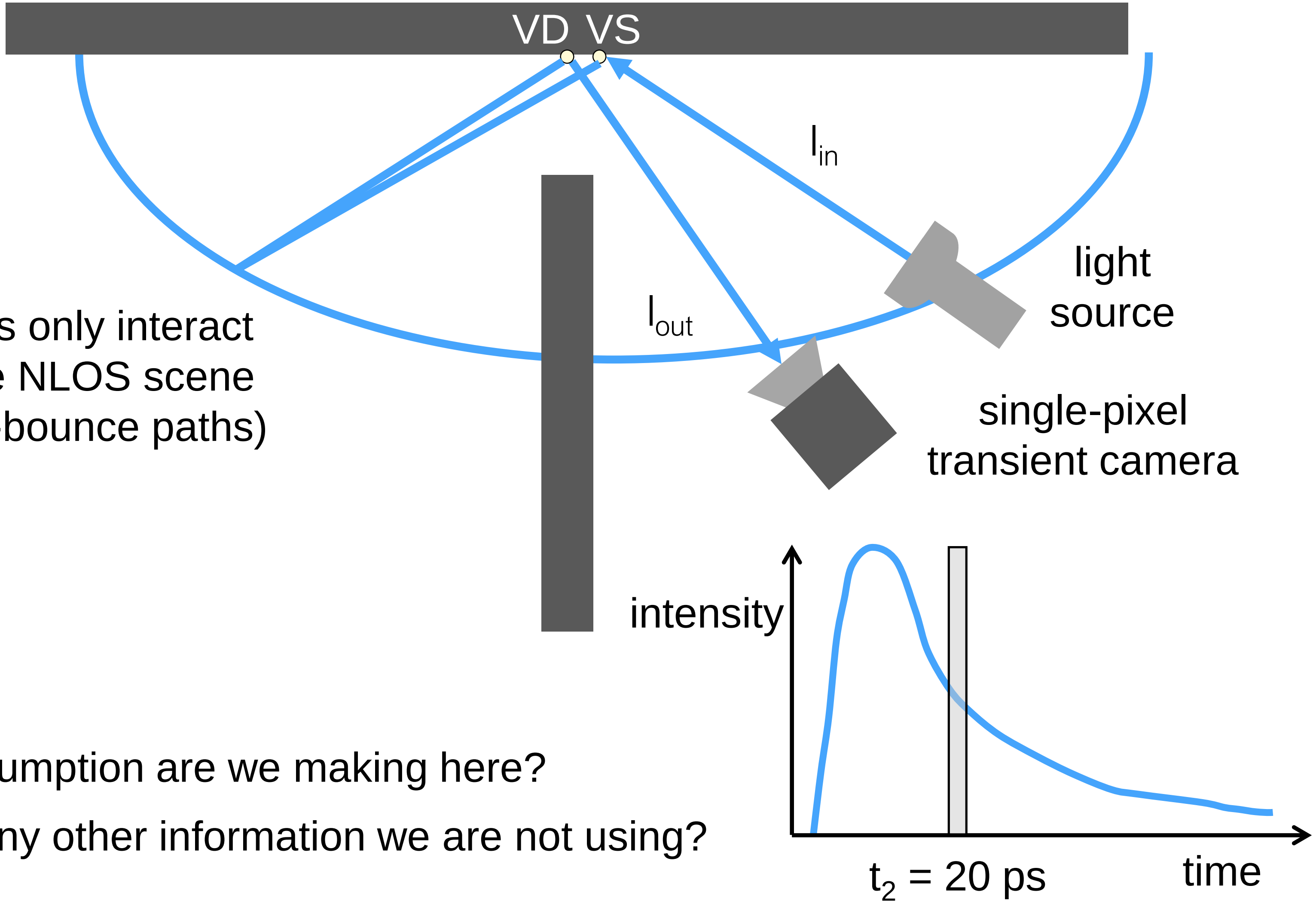
$$l_{nlos} = t_2 \cdot c - l_{in} - l_{out}$$

Processing a single-pixel transient



- What assumption are we making here?
- Is there any other information we are not using?

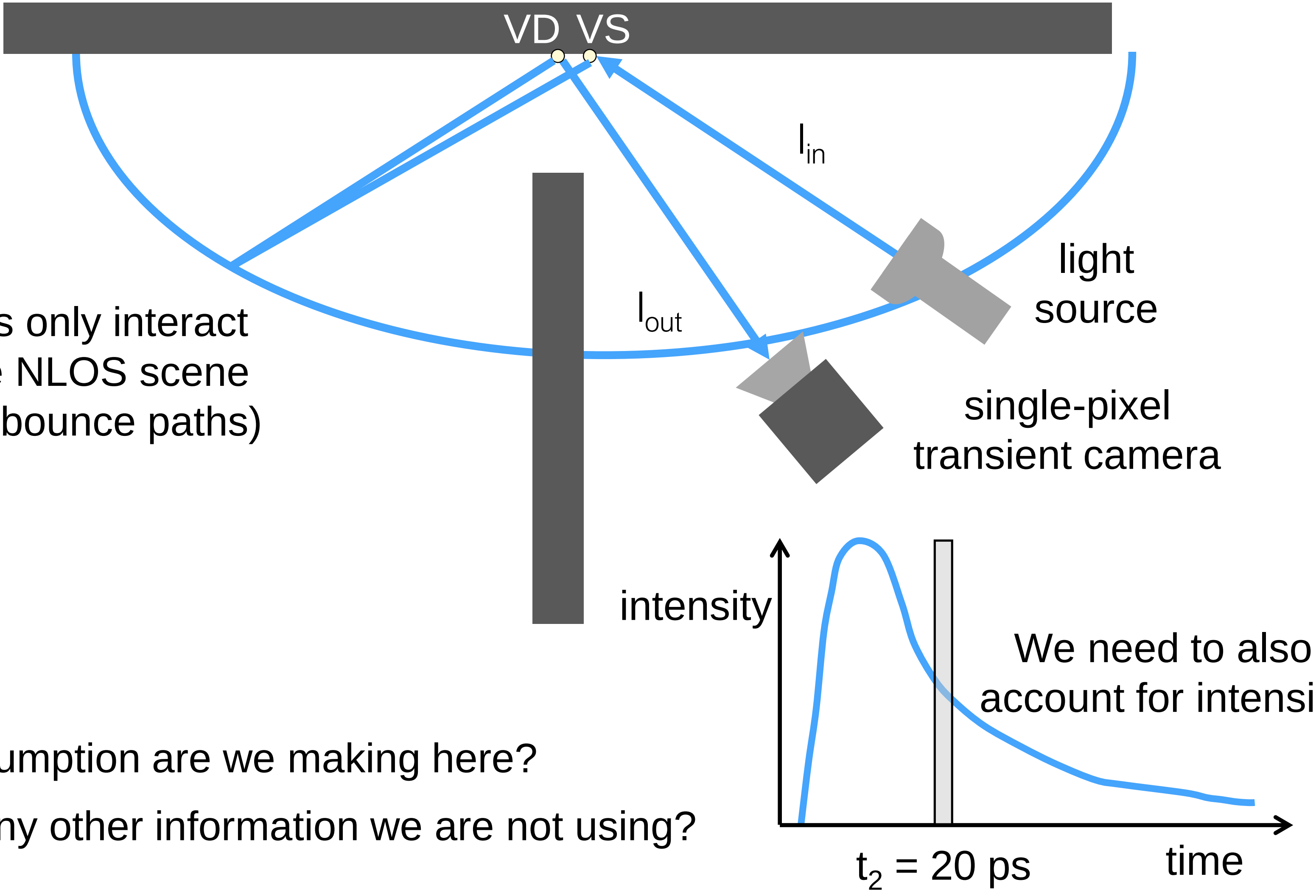
Processing another single-pixel transient



Photons only interact with the NLOS scene once (3-bounce paths)

- What assumption are we making here?
- Is there any other information we are not using?

Processing another single-pixel transient

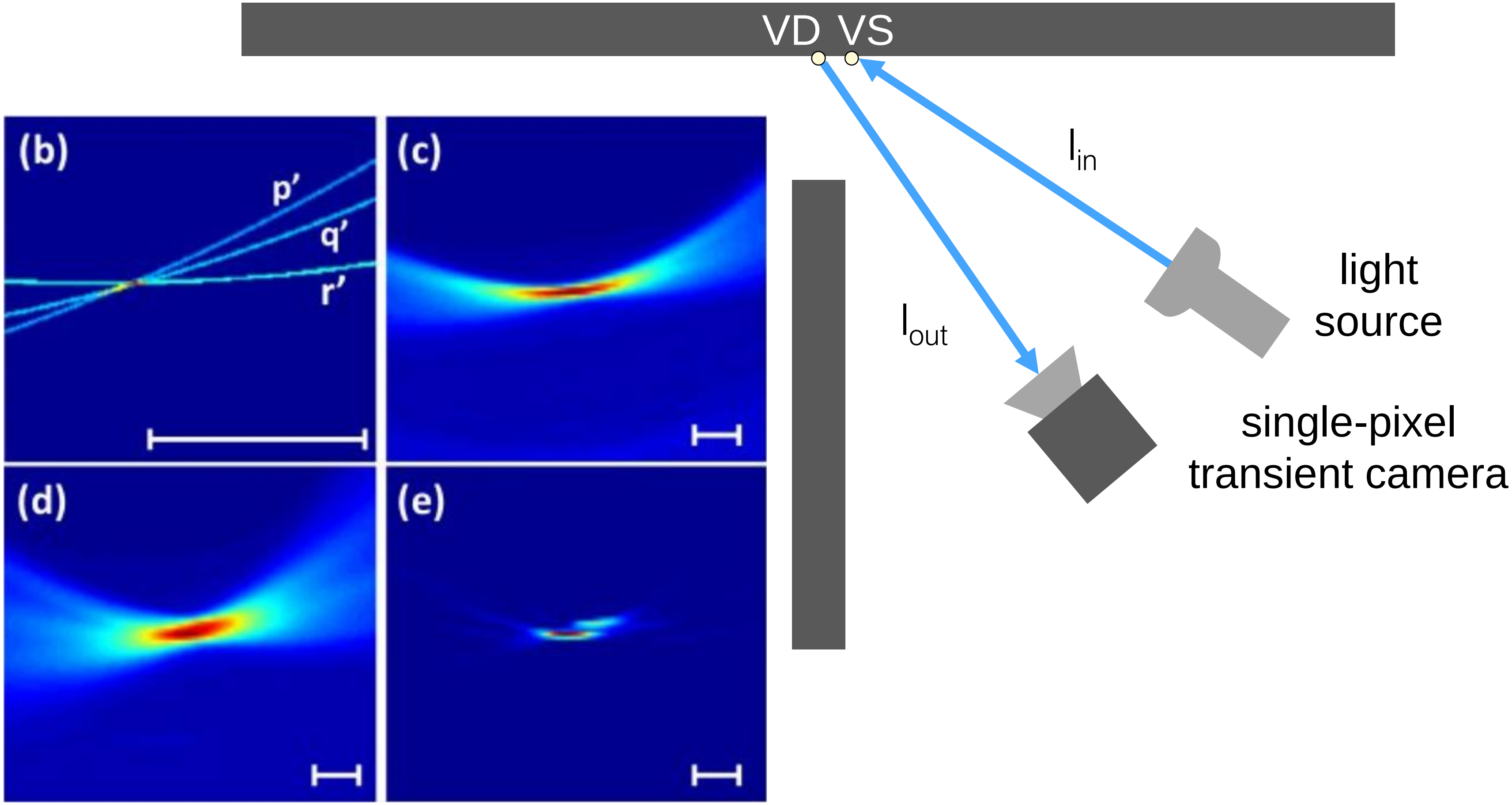


Photons only interact with the NLOS scene once (3-bounce paths)

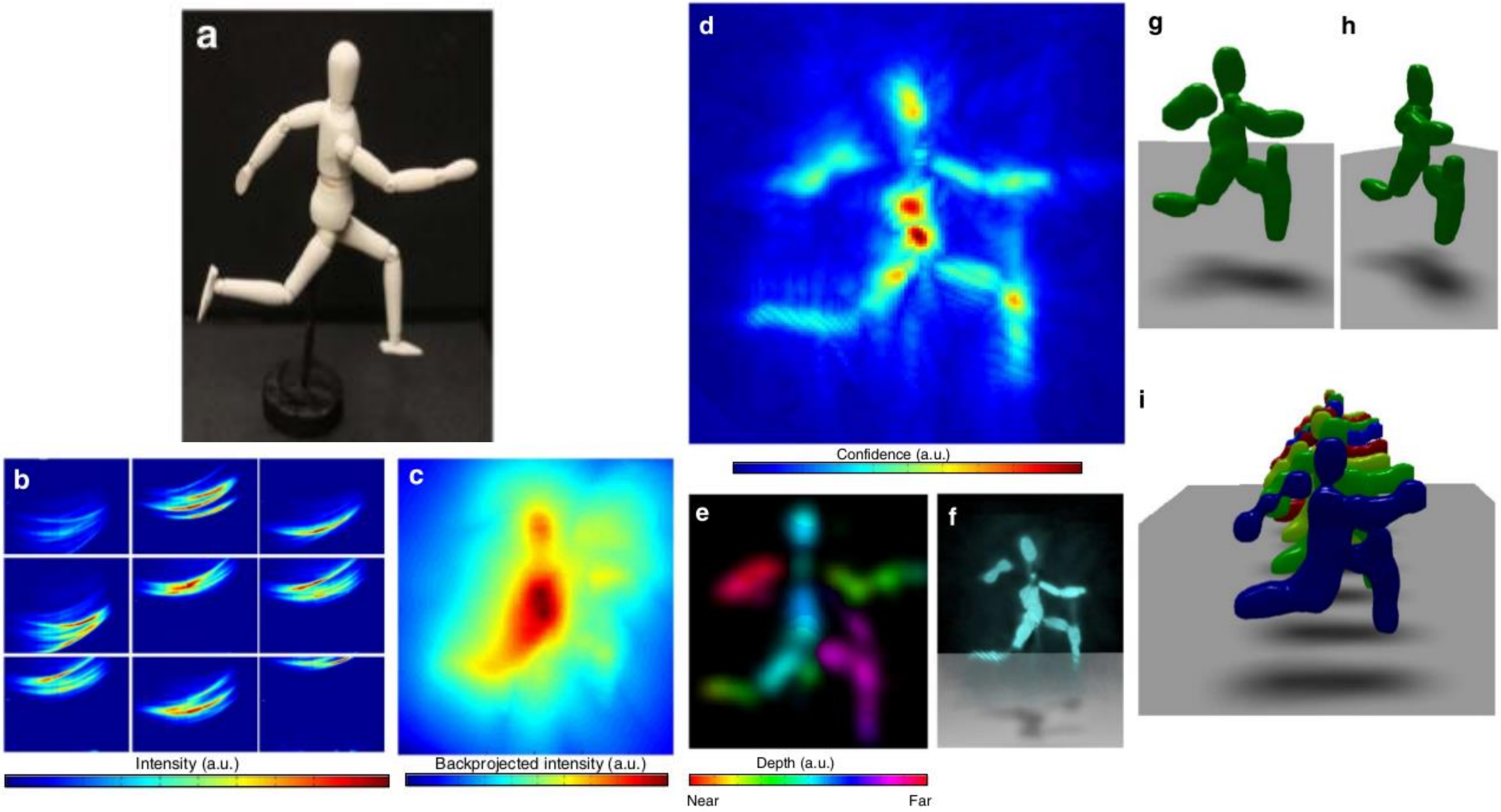
We need to also account for intensity

- What assumption are we making here?
- Is there any other information we are not using?

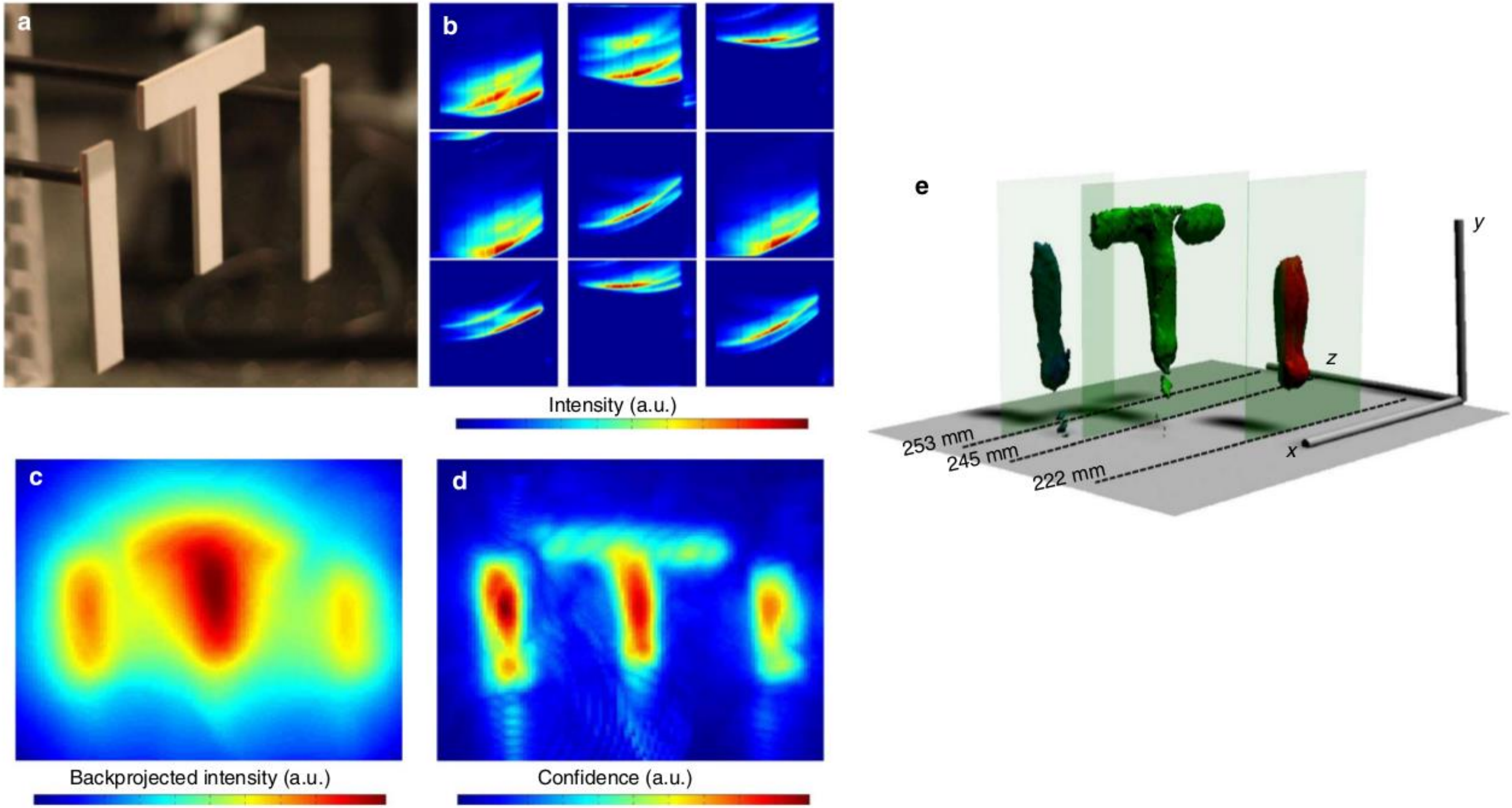
Elliptic backprojection



Elliptic backprojection



Elliptic backprojection

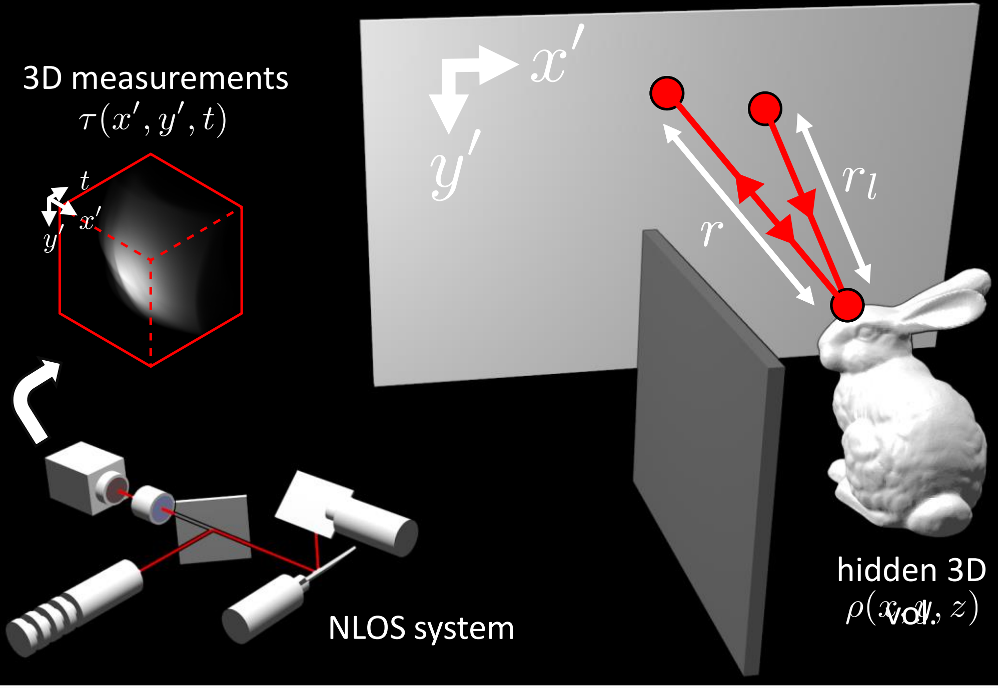


light cone transform

Our approach

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

1. confocalizing measurements
2. performing a change of variables (set $z = \sqrt{u}$, $t = 2\sqrt{v}/c$)



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

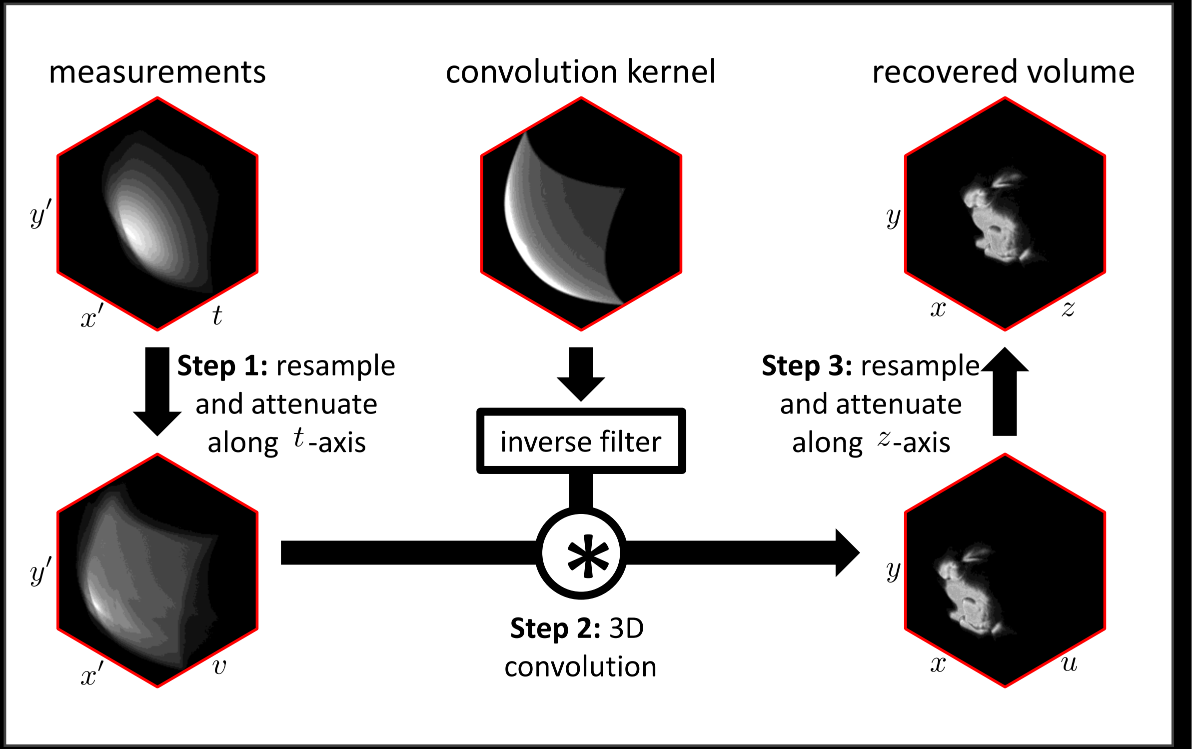
3D measurements radiometric term geometric term hidden 3D volume

$\tau = \mathbf{A}\,\alpha \times \rho * \rho$

(reparameterized NLOS image formation)

(convolutional NLOS image formation model)

light cone transform



measurements

convolution kernel

recovered volume

Step 1: resample and attenuate along t-axis

inverse filter

Step 3: resample and attenuate along z-axis

Step 2: 3D convolution

Alg. Complexity:

compute: $O(n^3 log(n))$
memory: $O(n^3)$

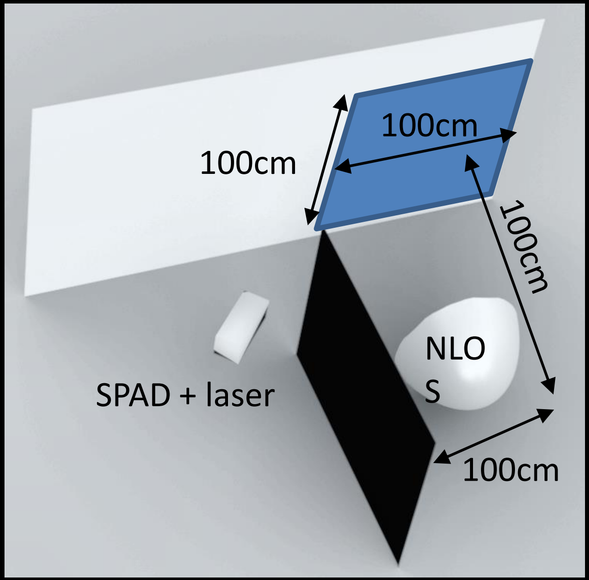
Current Implementation:

volume: 32 (x) x 32 (y) x 512 (t)
reconstruction rate: ~60 Hz
scan rate: ~2 Hz

Assumptions:

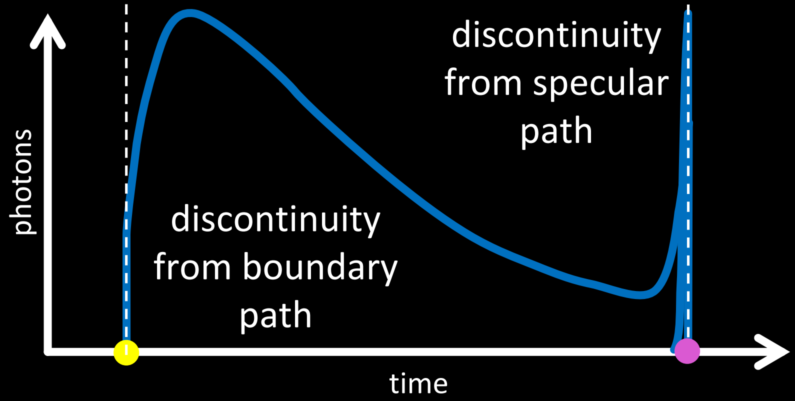
confocal scan
isotropic or retro. reflector
planar wall

Fermat paths: reconstruction based on geometric optics



setup

concave spherical shell
(Lambertian BRDF)



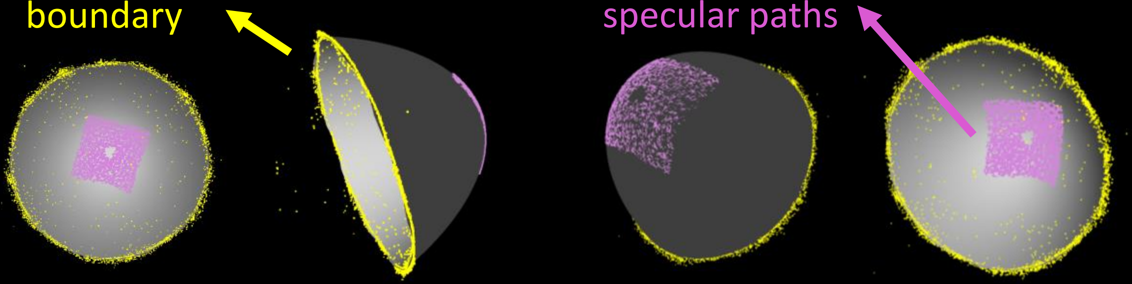
transient

We can uniquely reconstruct the NLOS surface point and normal as

$$x_{NLOS} = x_{LOS} - t_d(x_{LOS}) \cdot \nabla t_d(x_{LOS})$$

points at object boundary

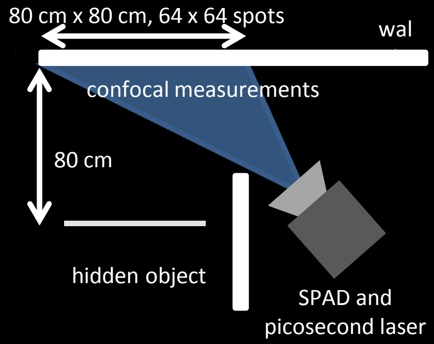
points forming specular paths



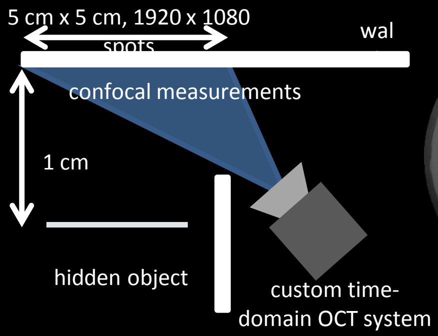
reconstructions shown in pink and yellow

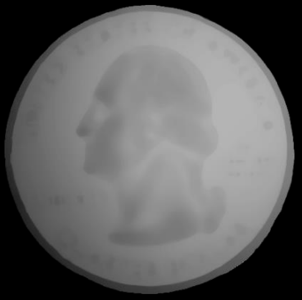
Surface reconstruction examples

shape around the corner with picosecond setup

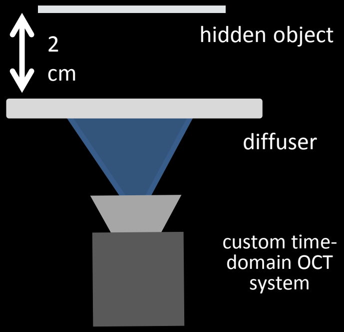


shape around the corner with femtosecond setup

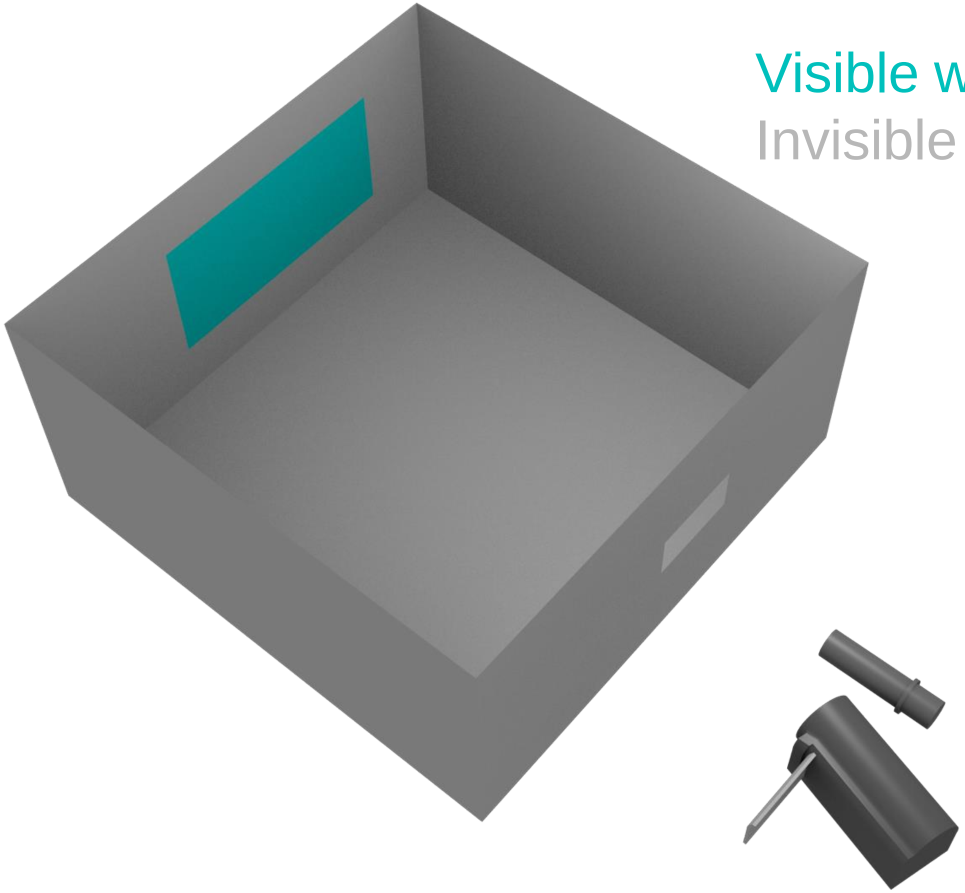




shape through diffuser with femtosecond setup



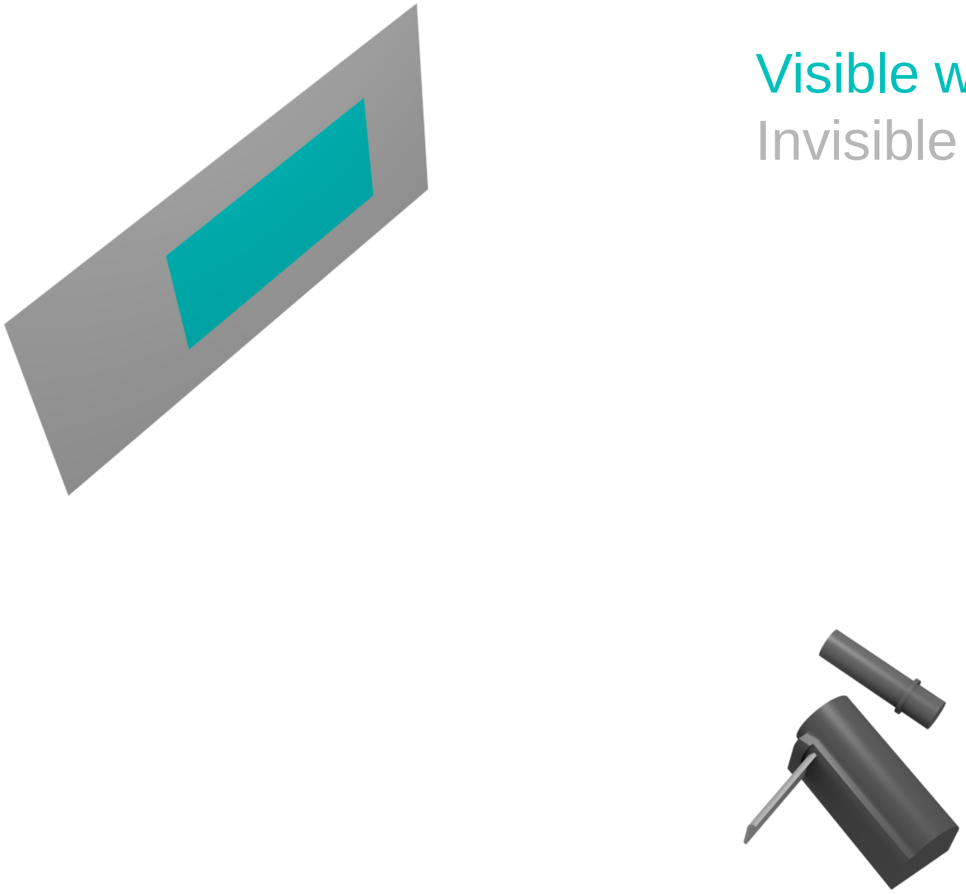
Reconstructing Hidden Rooms



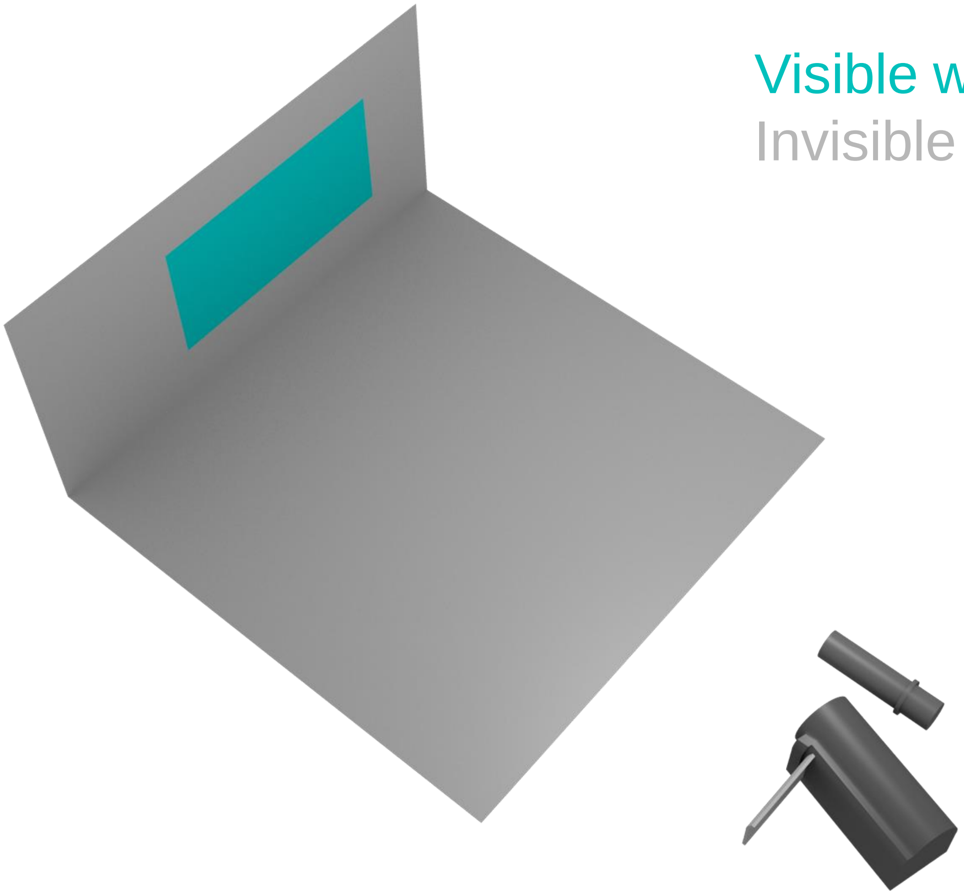
Visible wall

Invisible walls

Reconstructing Hidden Rooms: One plane at a time



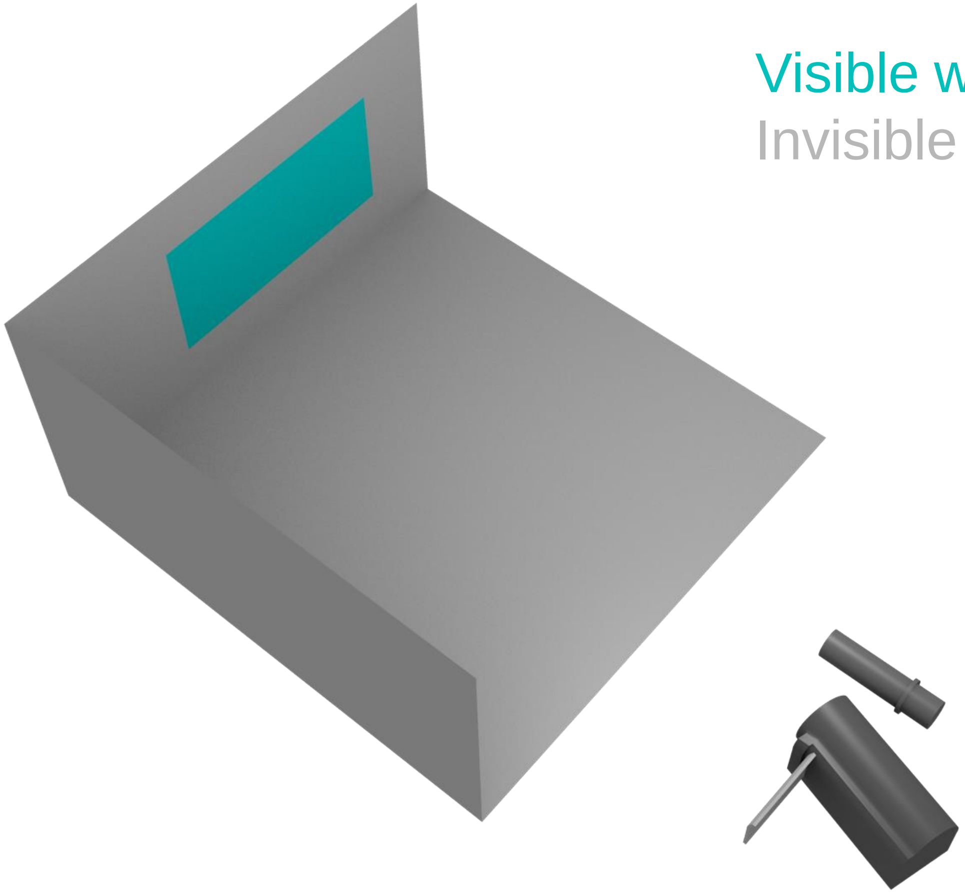
Reconstructing Hidden Rooms: One plane at a time



Visible wall

Invisible walls

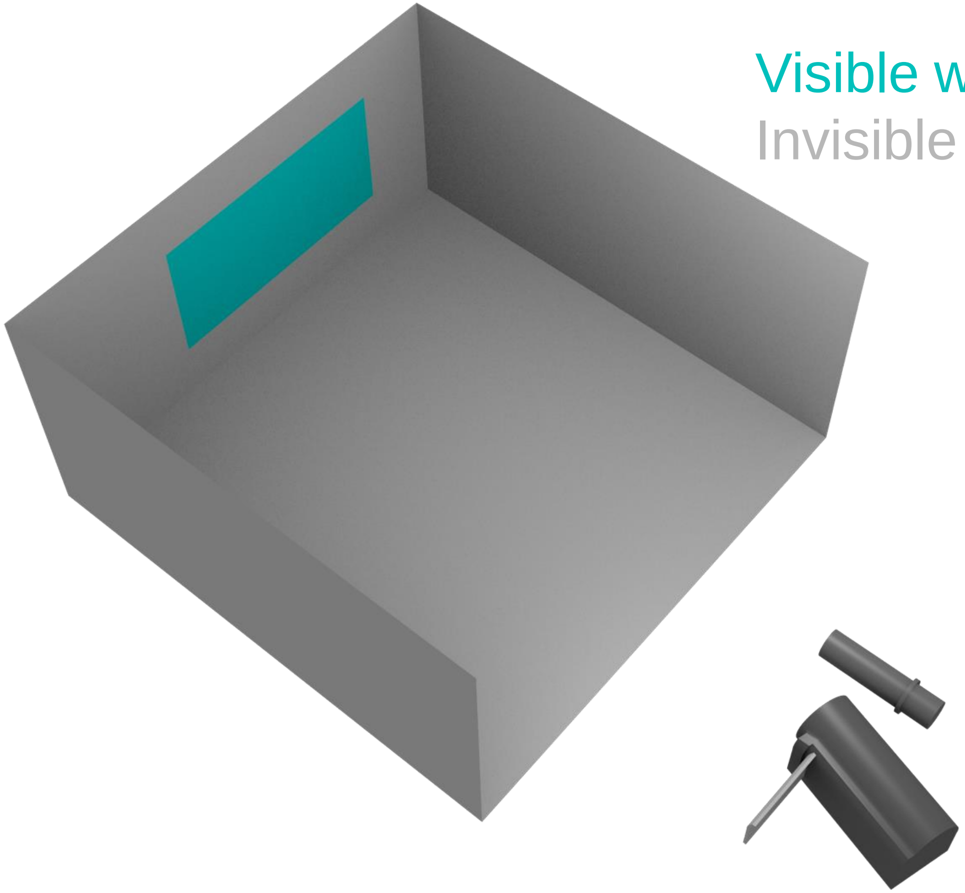
Reconstructing Hidden Rooms: One plane at a time



Visible wall

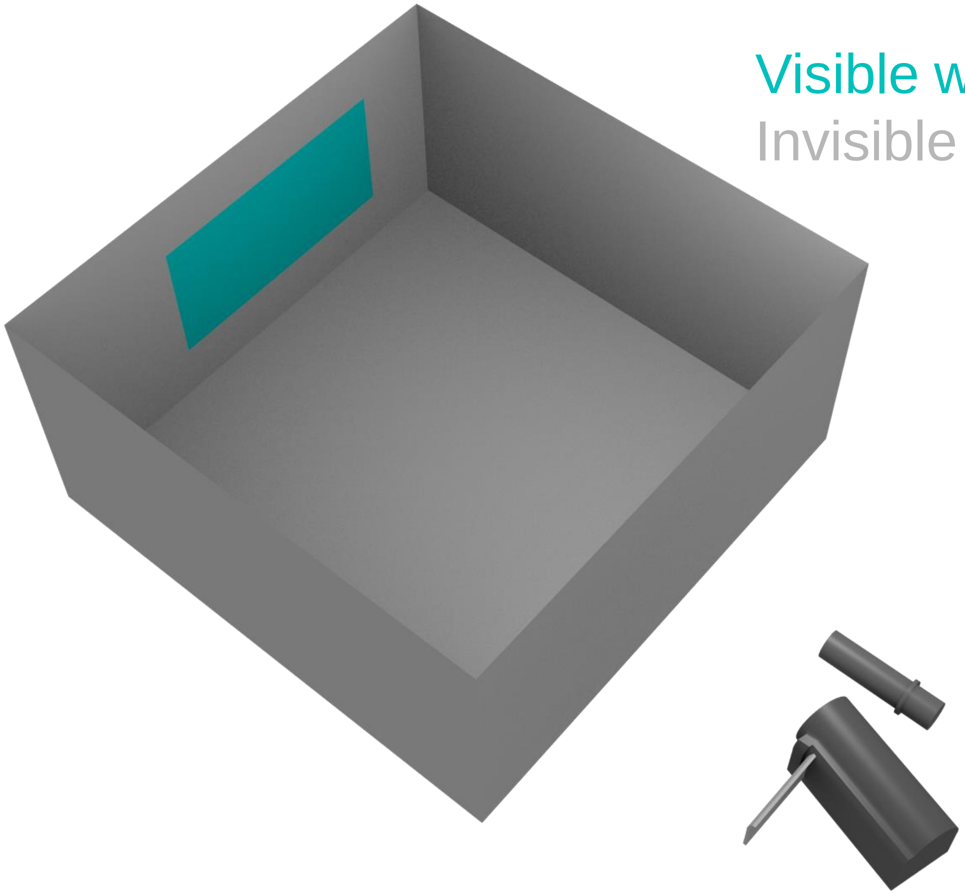
Invisible walls

Reconstructing Hidden Rooms: One plane at a time



Visible wall
Invisible walls

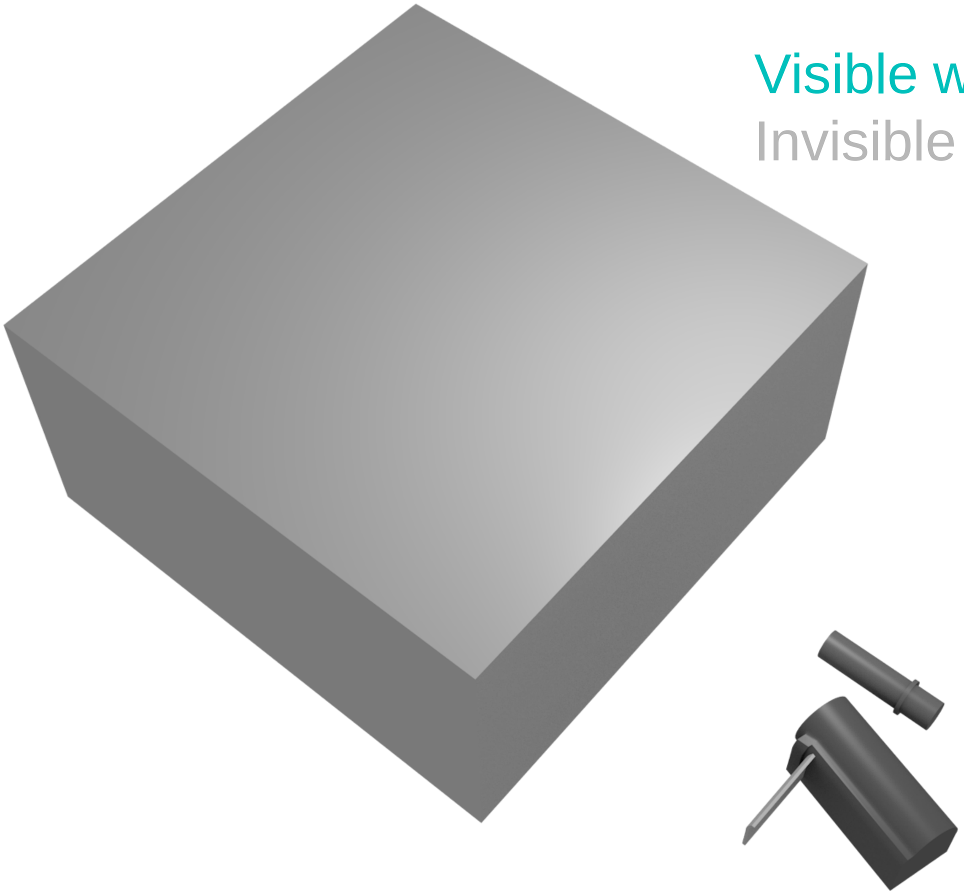
Reconstructing Hidden Rooms: One plane at a time



Visible wall

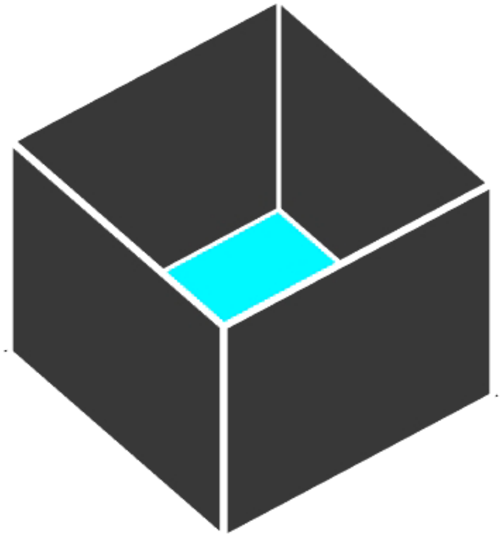
Invisible walls

Reconstructing Hidden Rooms: One plane at a time

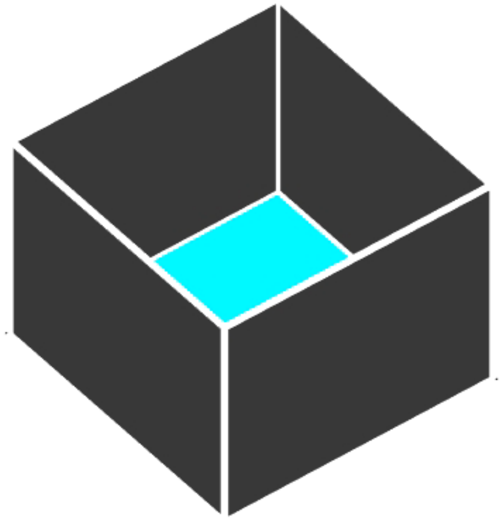


Visible wall
Invisible walls

Reconstructing Rectangular Rooms



Original room

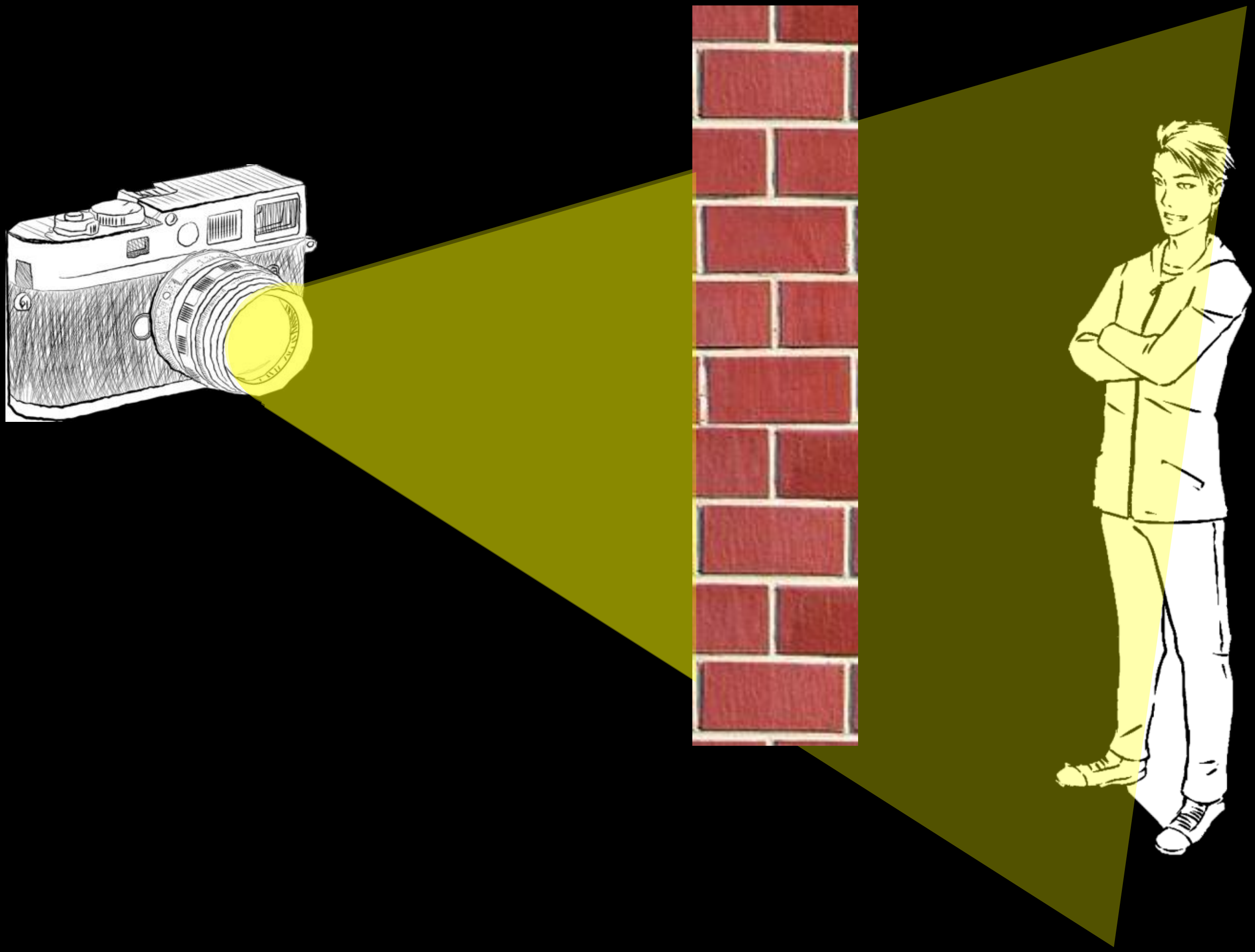


Reconstructed room

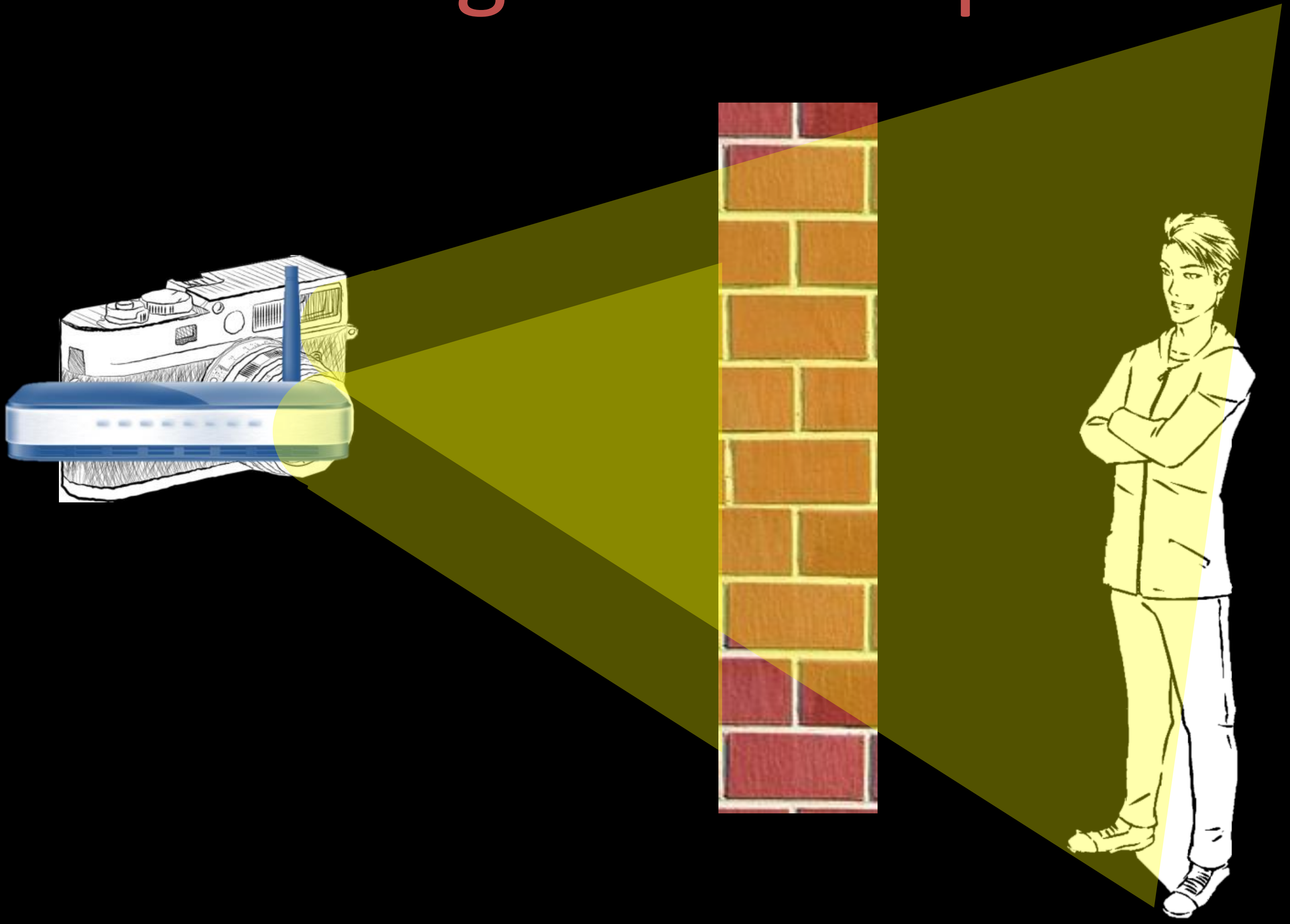
- Average AICP error for all the walls is TBD mm (TBD %). – Normalized with average room length of 1.1m

Active NLOS imaging using WiFi

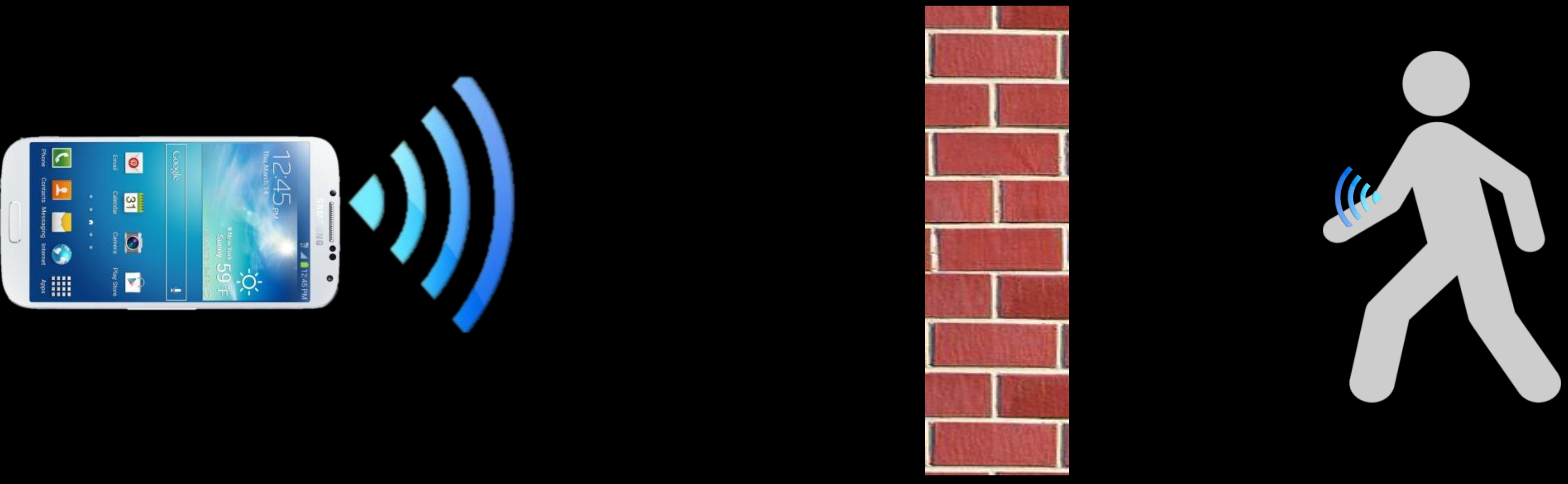
Imaging through occlusions



Imaging through occlusions
using radio frequencies



Key Idea



Challenges

Wall refection is 10,000x stronger than reflections coming from behind the wall

Wi-Vi: Small, Low-Power, Wi-Fi

- Eliminate the wall’s reflection
- Track people from reflections
- Gesture-based interface
- Implemented on software radios

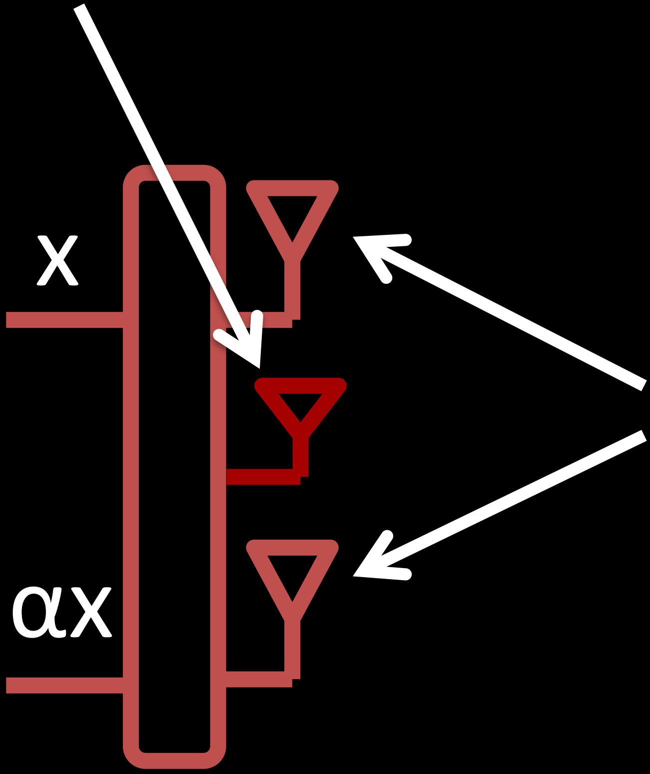
How Can We Eliminate the Wall’s Reflection?

Idea: transmit two waves that cancel each other when they reflect off static objects but not moving objects

Wall is static → disappears

People tend to move → detectable

Eliminating the Wall’s Reflection

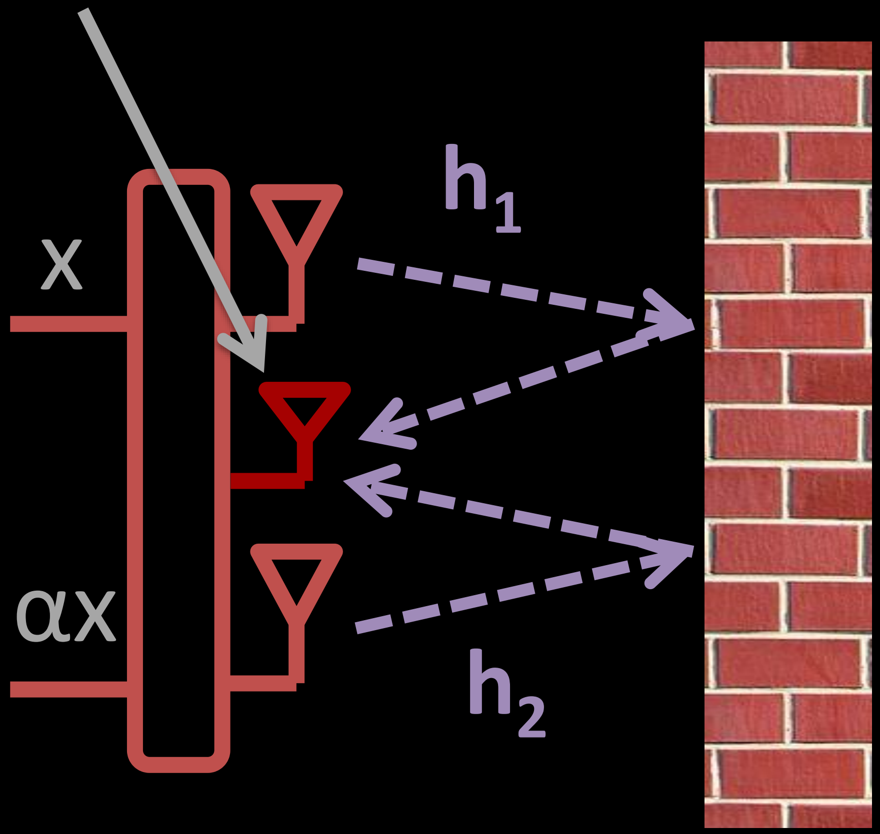


Receive Antenna:

Transmit Antennas

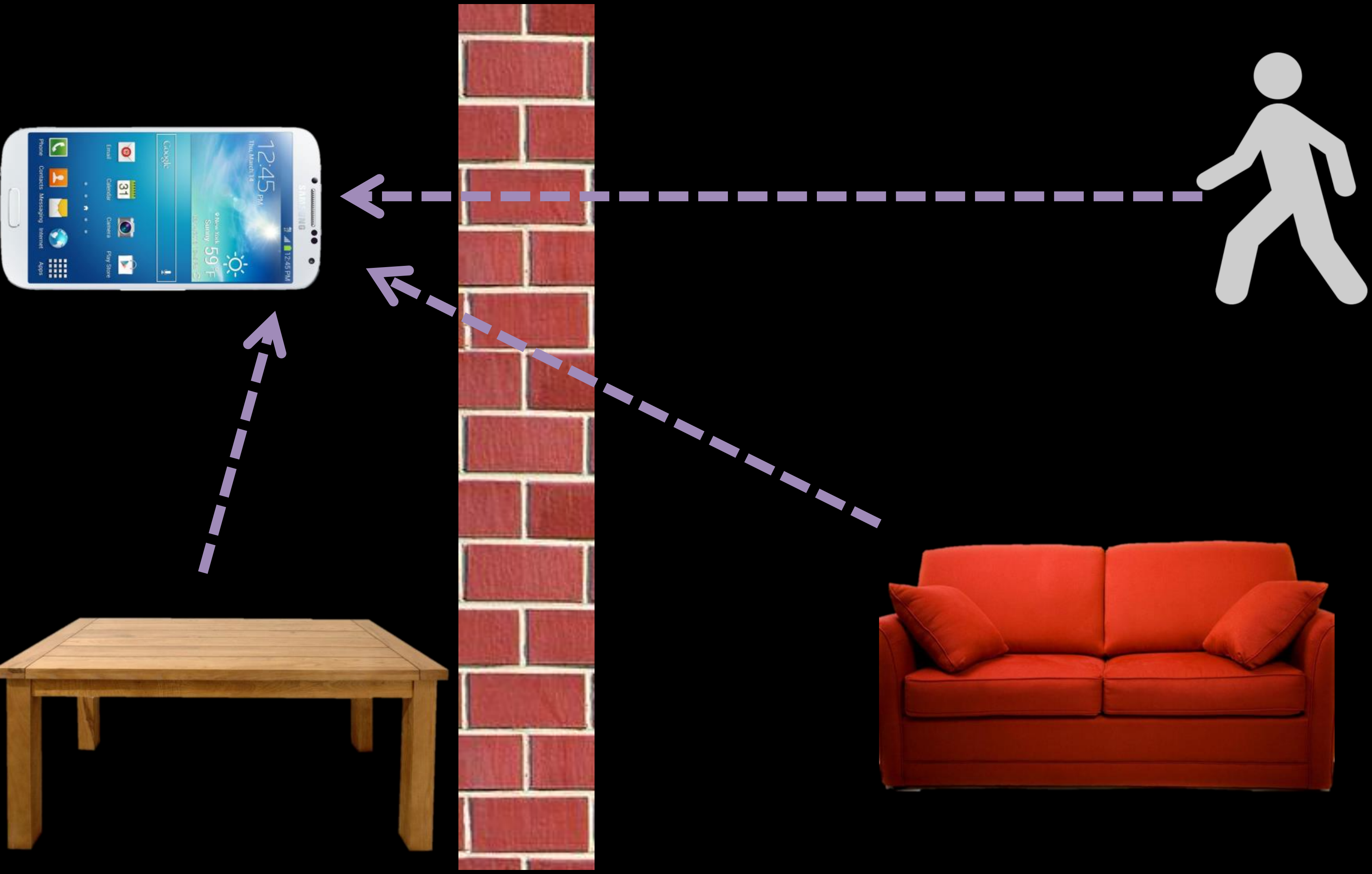
Eliminating the Wall’s Reflection

Receive Antenna: $y = h_1 x + h_2 \alpha x \longrightarrow 0$

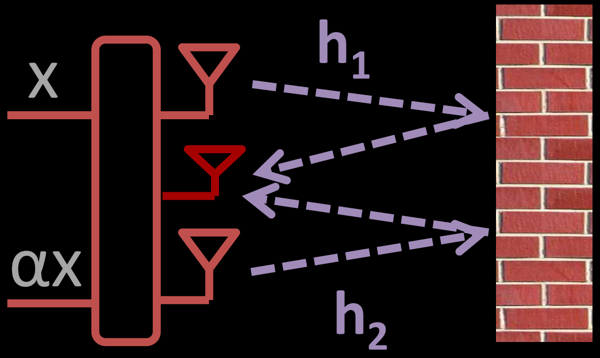


$\alpha = -h_1 / h_2$

Eliminating All Static Reflections



Eliminating All Static Reflections



$$y = h_1 x + h_2 \alpha x$$

Static objects (wall, furniture, etc.) have constant channels

$$y = \cancel{h_1 x + h_2(- h_1/ h_2)x} \rightarrow 0$$

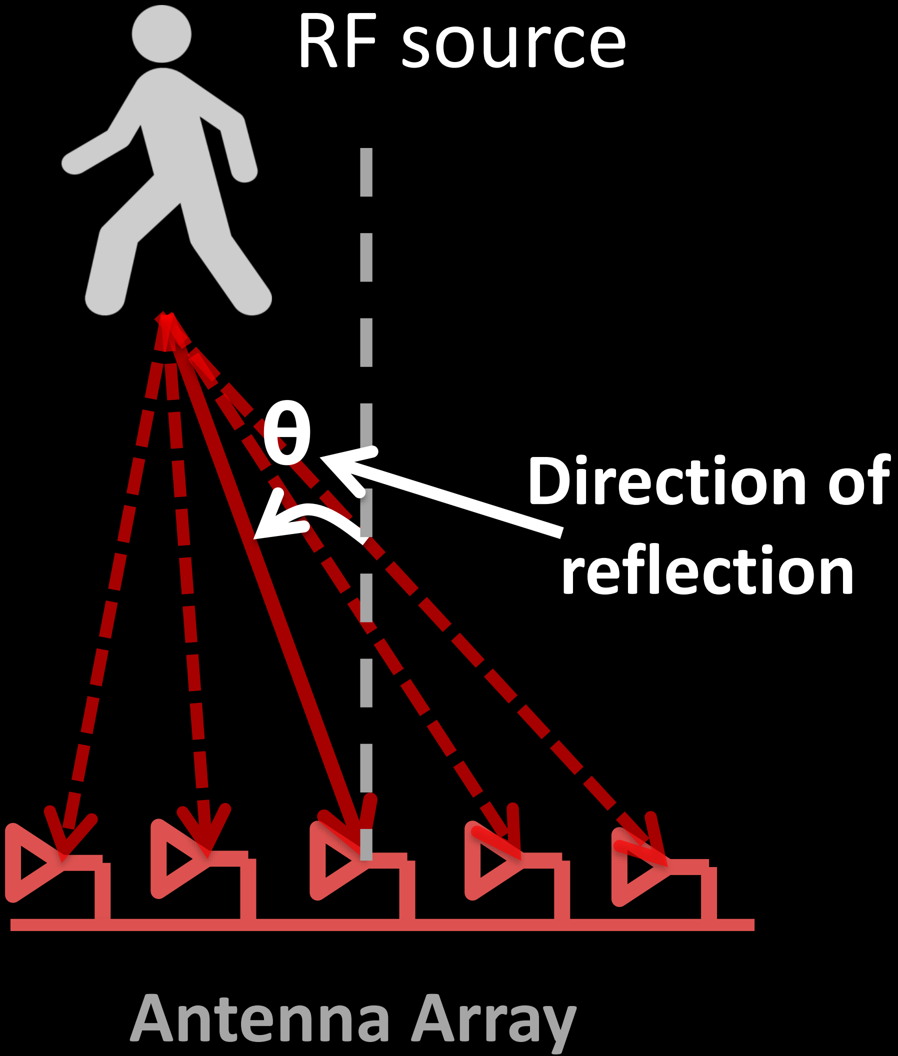
People move, therefore their channels change

$$y = h_1' x + h_2'(- h_1/ h_2)x$$

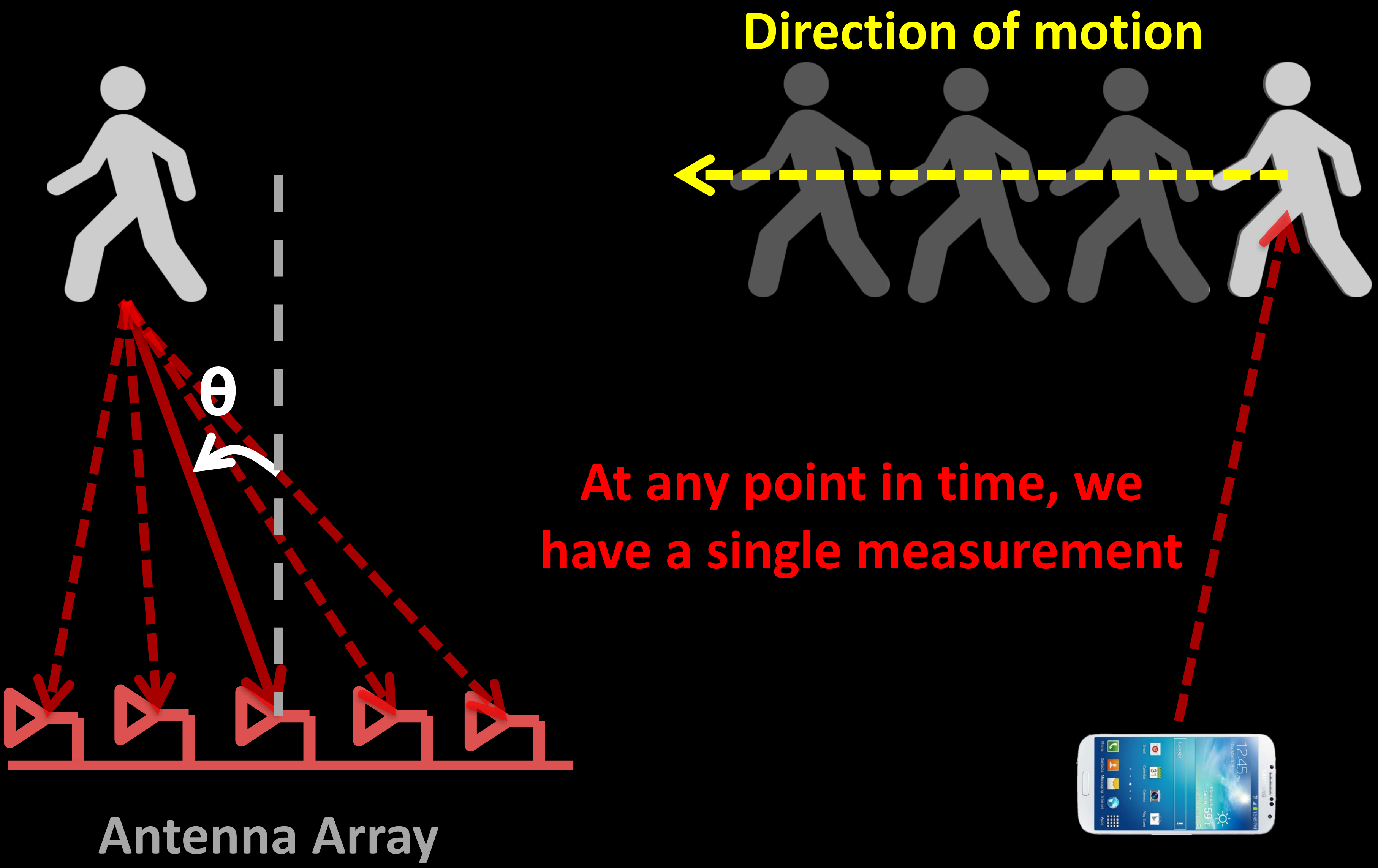
Not Zero

How Can We Track Using Reflections?

Tracking Motion

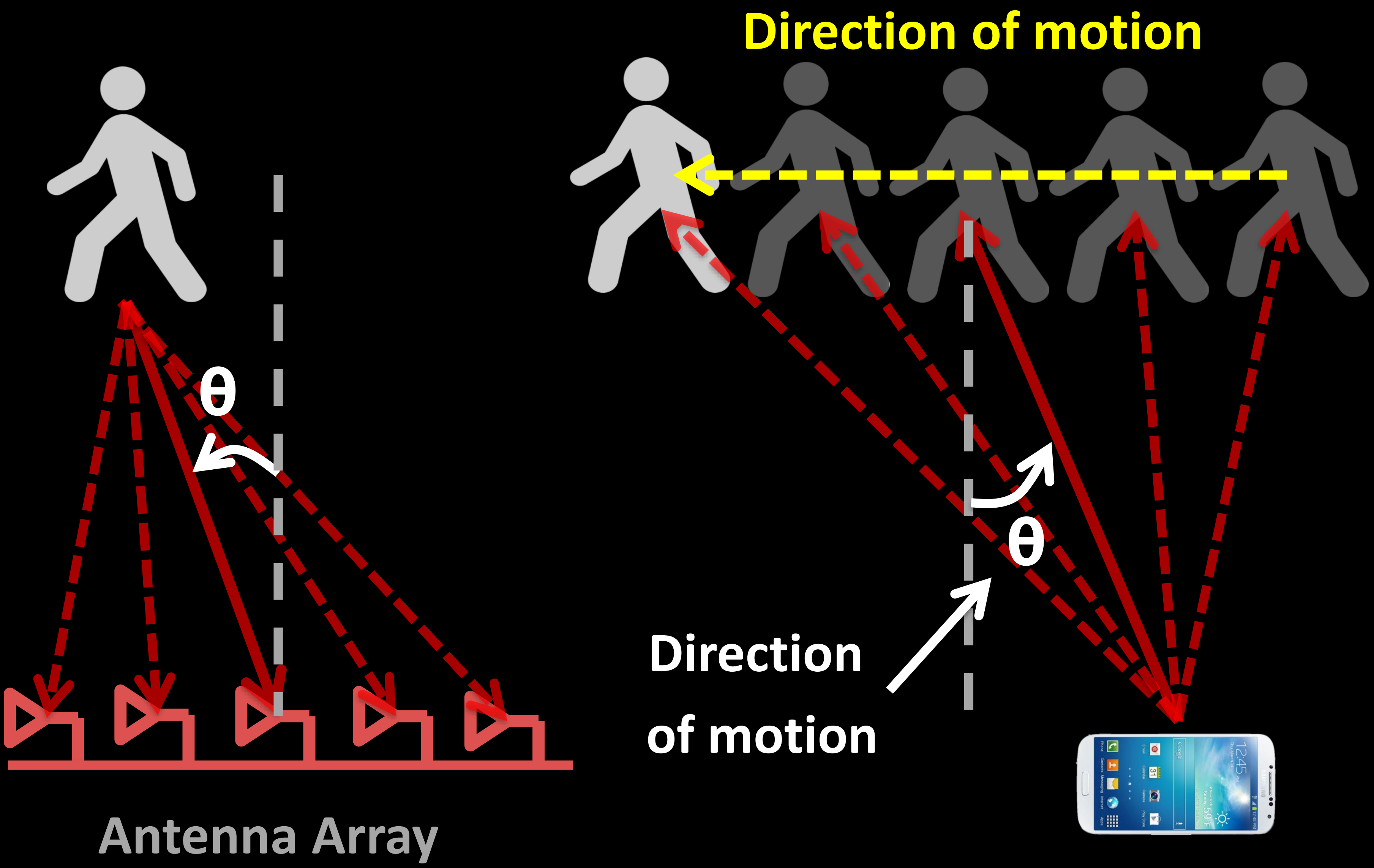


Tracking Motion



At any point in time, we have a single measurement

Tracking Motion



Direction of motion

θ

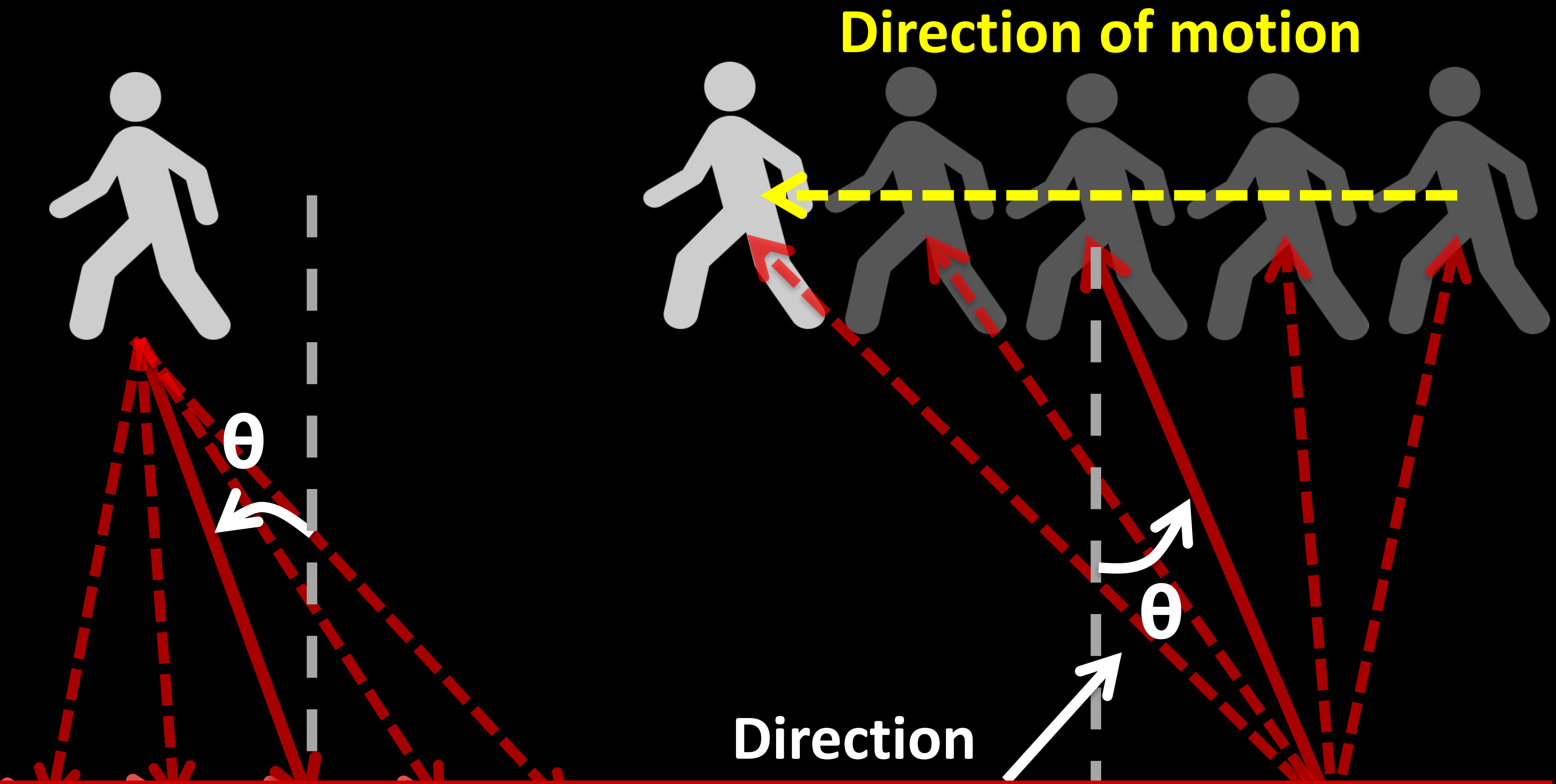
Antenna Array

θ

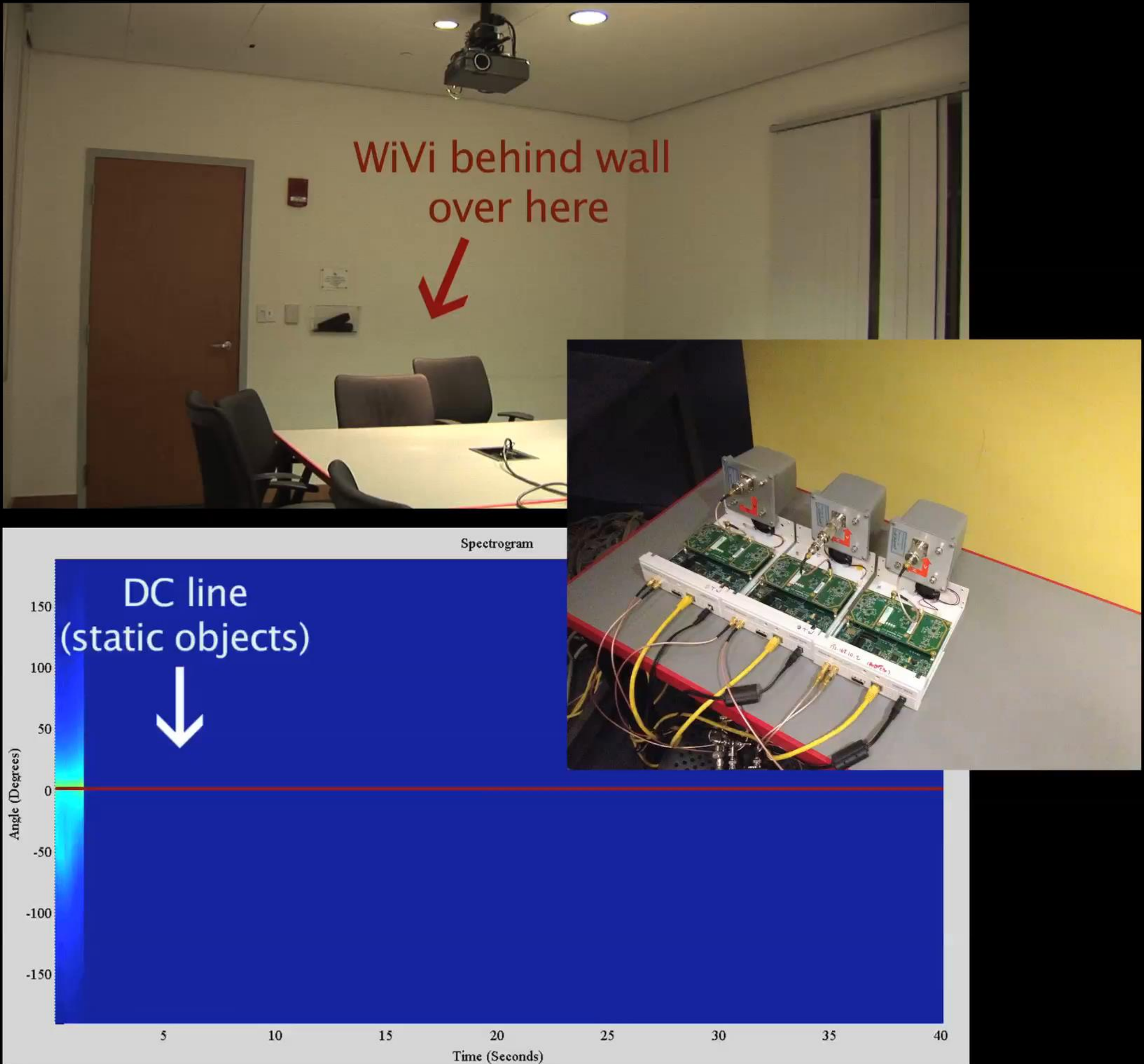
Direction

of motion

Tracking Motion



Human motion emulates antenna array



WiVi behind wall
over here

DC line
(static objects)

A Through-Wall Gesture Interface

- Sending Commands with Gestures

- Two simple gestures to represent bit ‘0’ and bit ‘1’

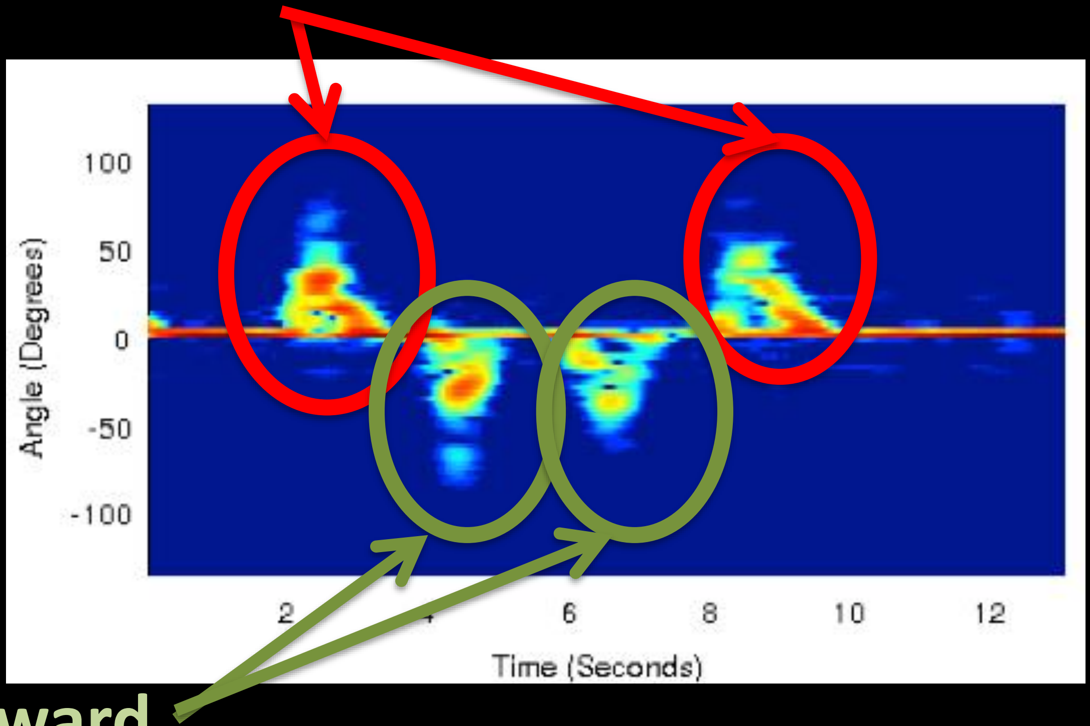
- Can combine sequence of gestures to convey longer message

Gesture Encoding

Bit ‘0’: step forward followed by step backward

Bit ‘1’: step backward followed by step forward

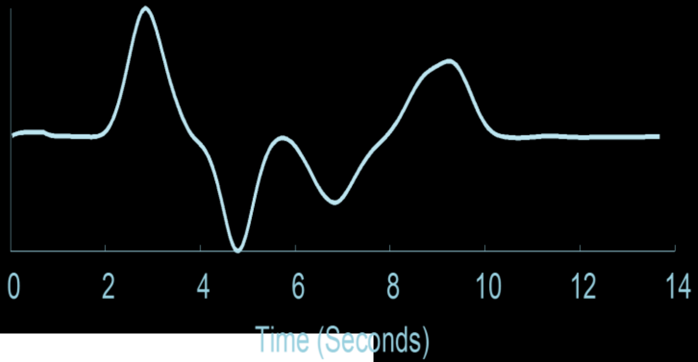
Step Forward



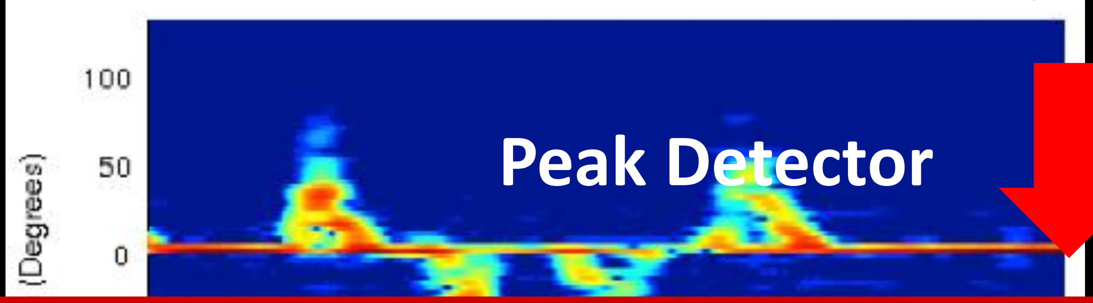
Step Backward

Gesture Decoding

Matched Filter

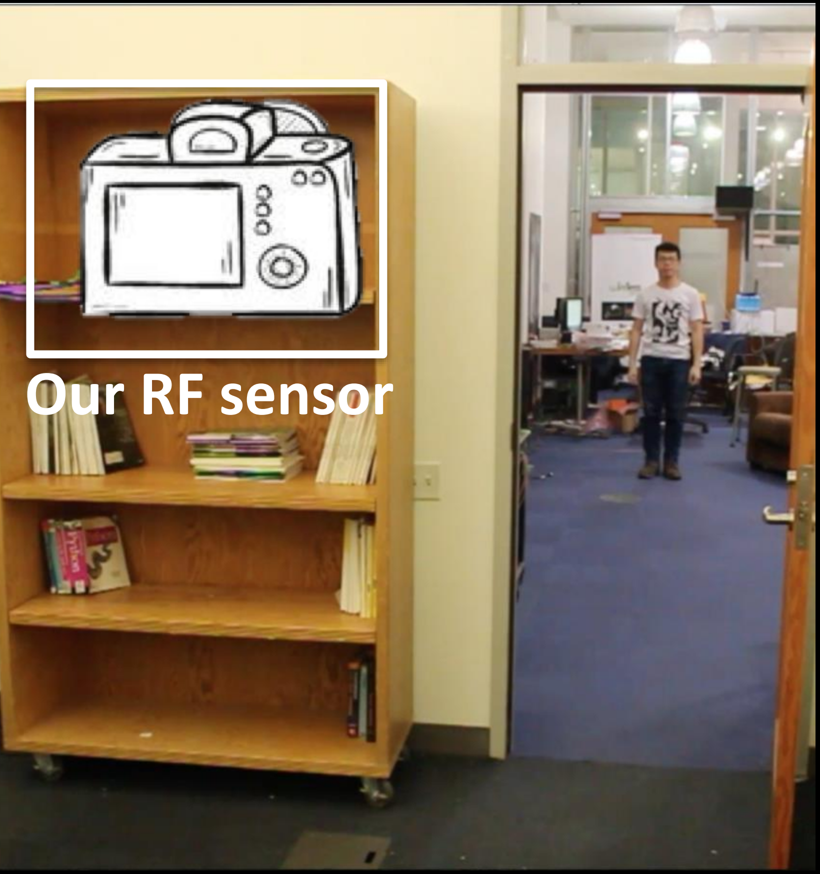


Peak Detector



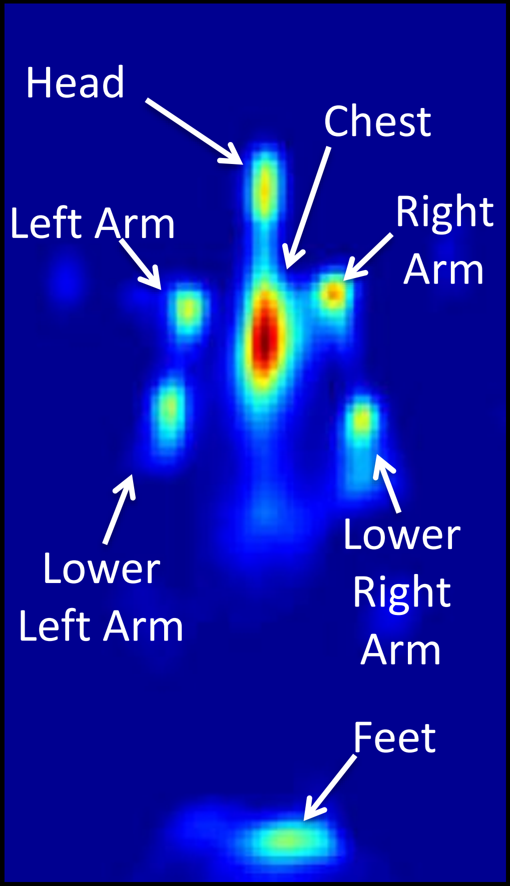
Gesture interface that works through walls and none-line-of-sight

Imaging through occlusions
using radio frequencies



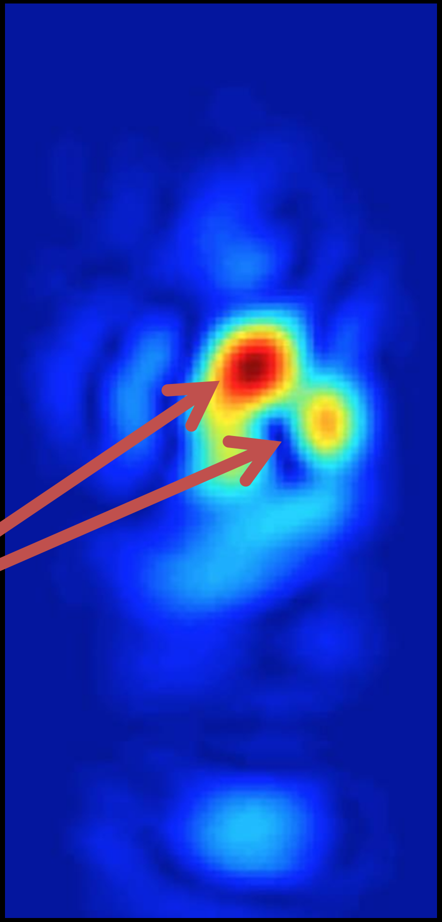
Camera image

Our output



Challenge: Don't get reflections from most points in RF

Output of 3D RF Scan



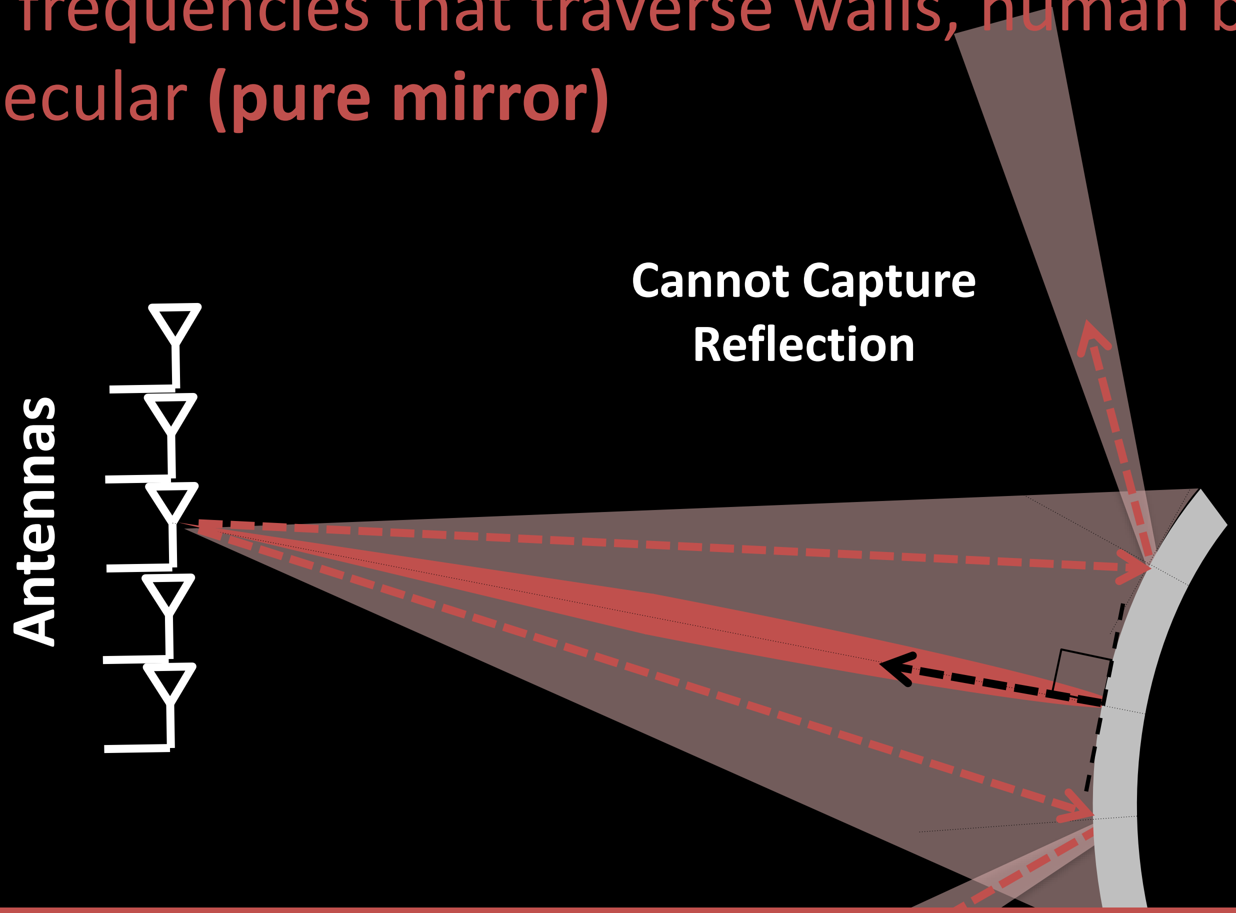
Blobs of reflection power

Challenge: Don’t get reflections from most points in RF

At frequencies that traverse walls, human body parts are specular **(pure mirror)**

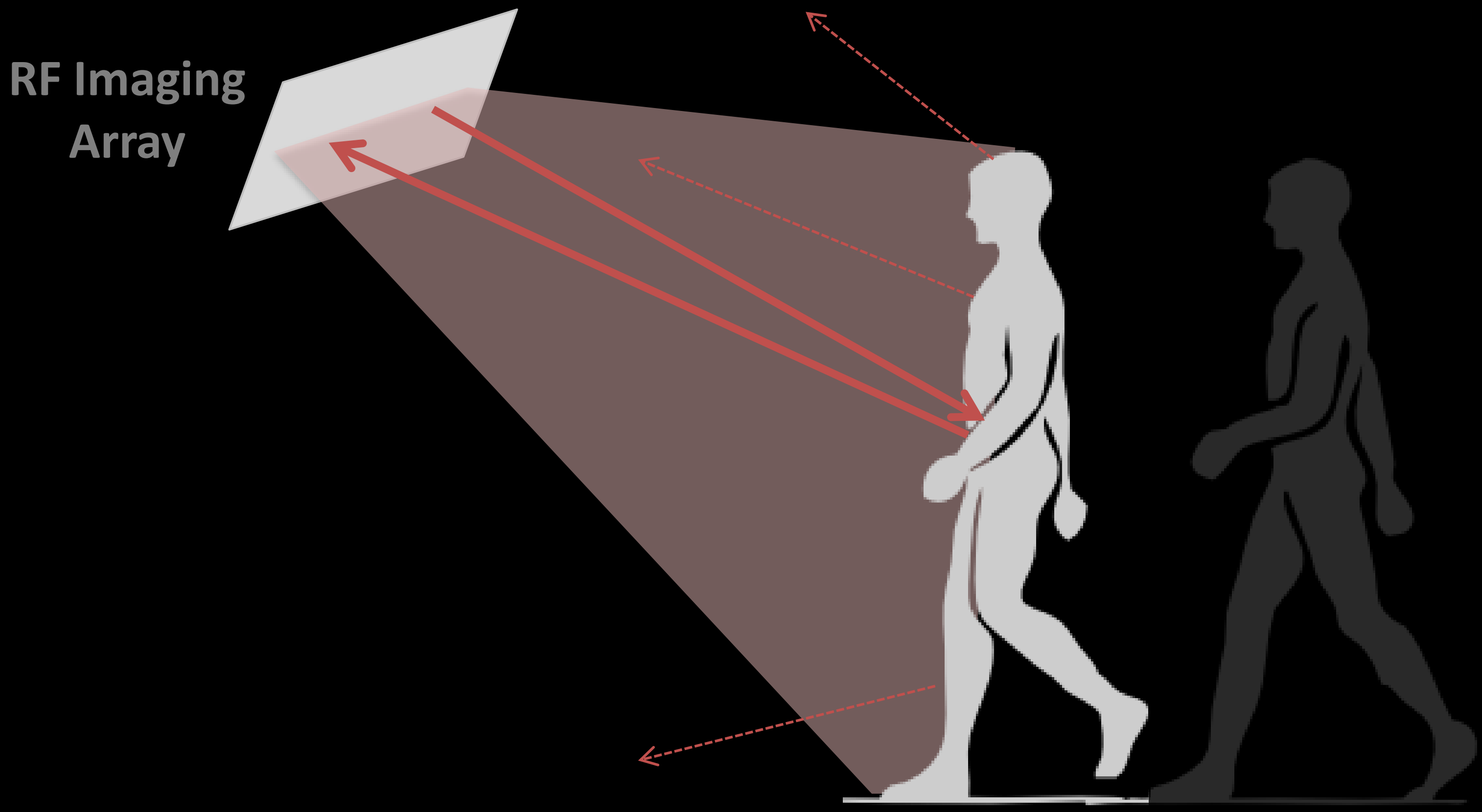
Challenge: Don't get reflections from most points in RF

At frequencies that traverse walls, human body parts are specular **(pure mirror)**

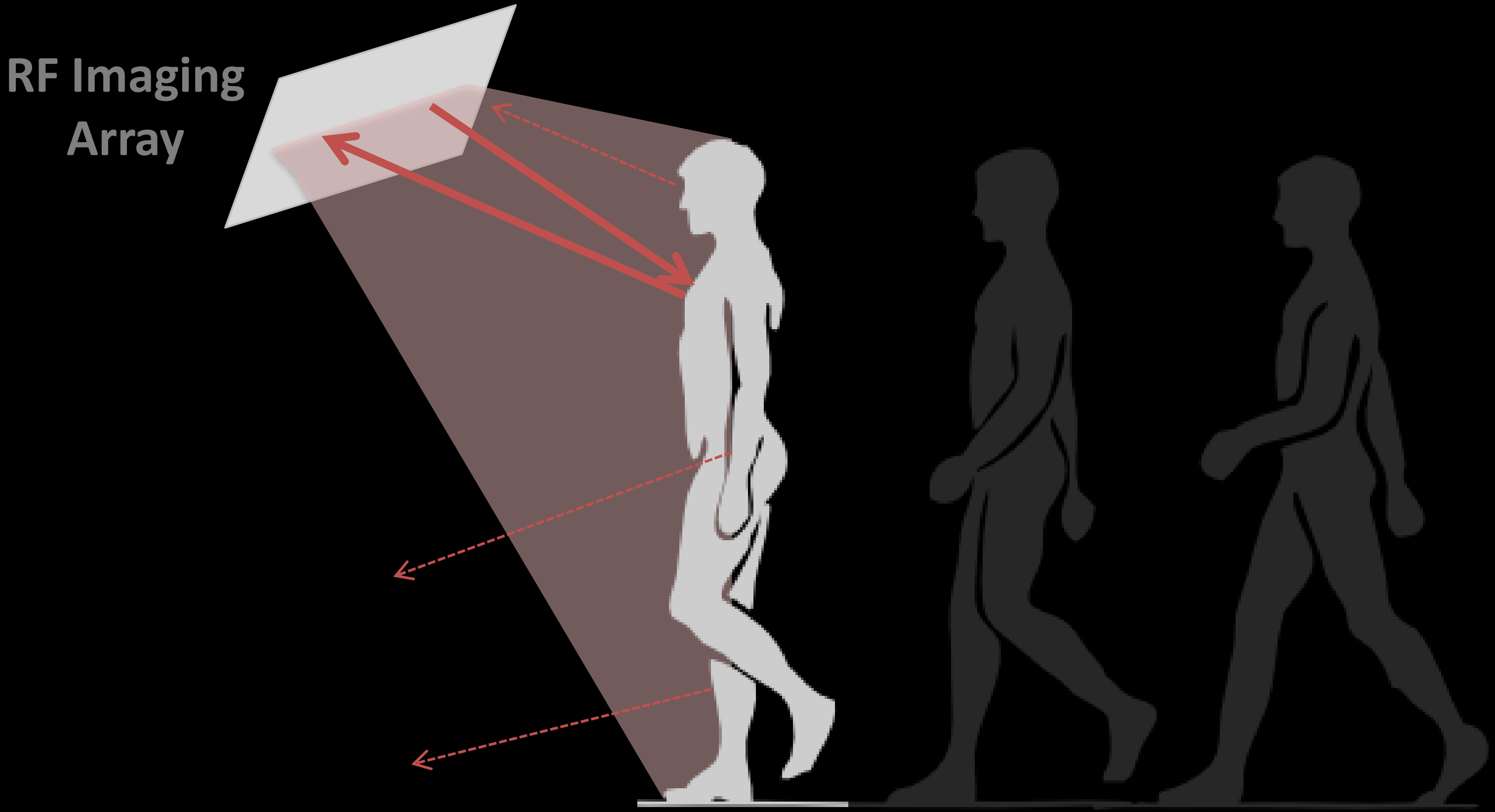


At every point in time, get reflections from only a subset of body parts

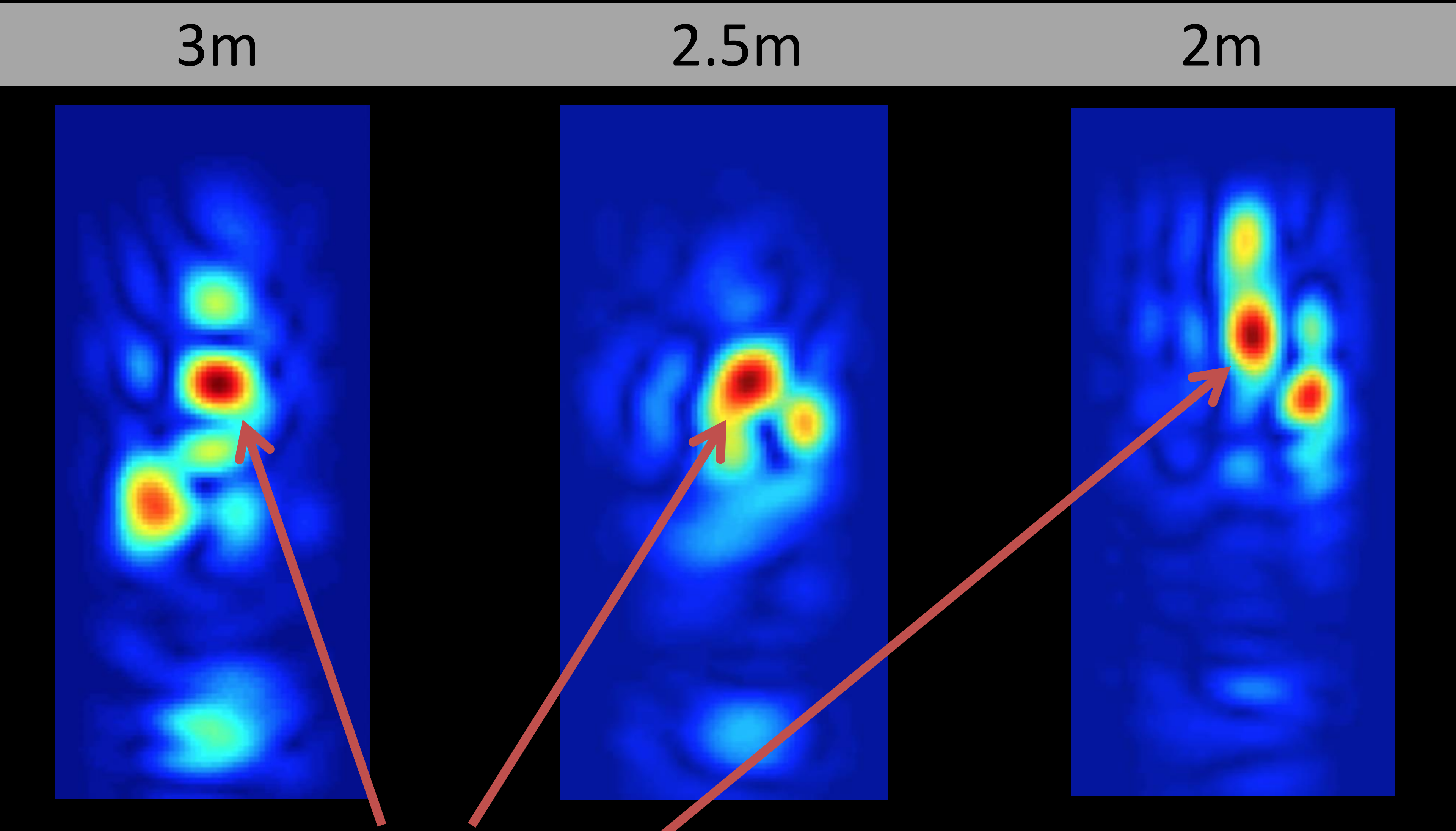
Solution Idea: Exploit Human Motion and Aggregate over Time



Solution Idea: Exploit Human Motion and Aggregate over Time

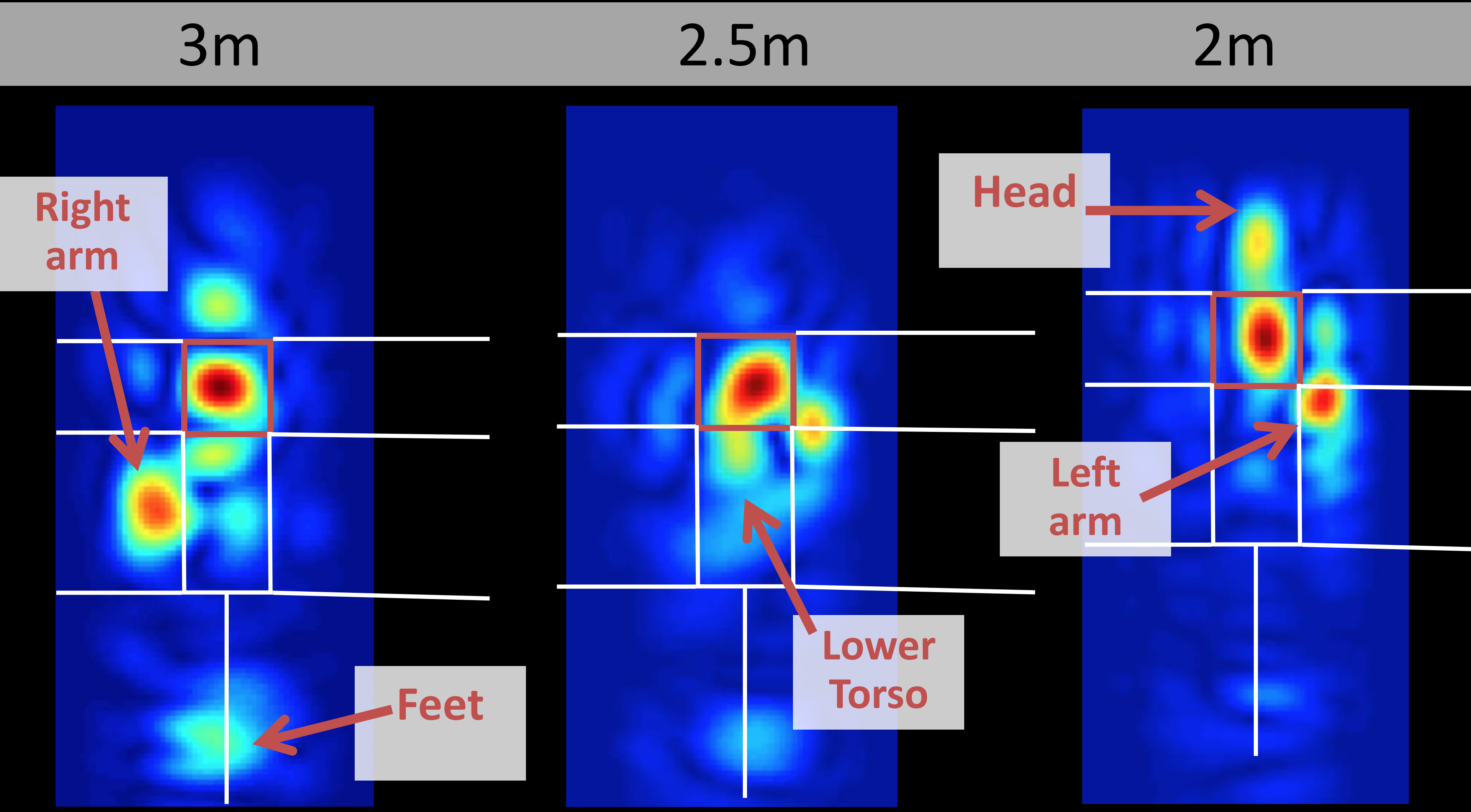


Human Walks toward Sensor



Chest (Largest Convex Reflector) → Use it as a pivot: for motion compensation and segmentation

Human Walks toward Sensor

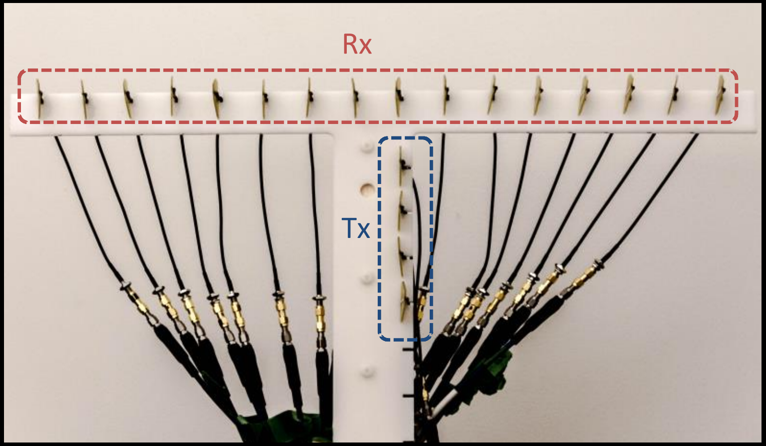


Combine the various snapshots

Human Walks toward Sensor

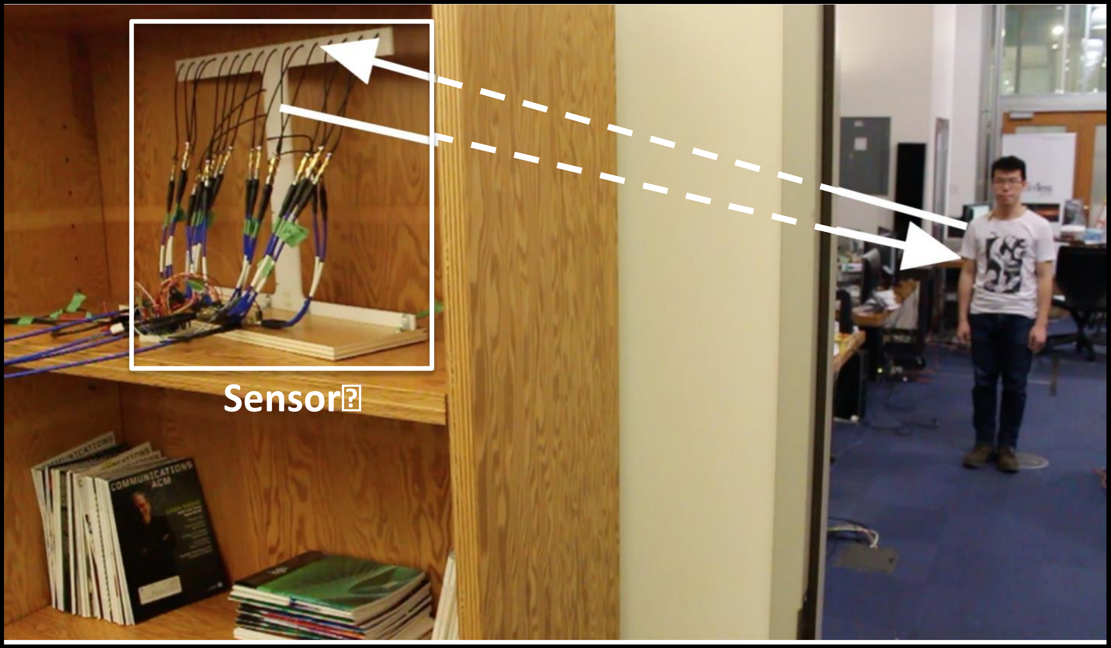
Implementation

- Hardware
  - 2D Antenna Array
  - Built RF circuit
    - 1/1,000 power of WiFi
    - USB connection to PC



- Software
  - Coarse-to-fine algorithm implemented in GPU to generate reflection snapshots in real-time

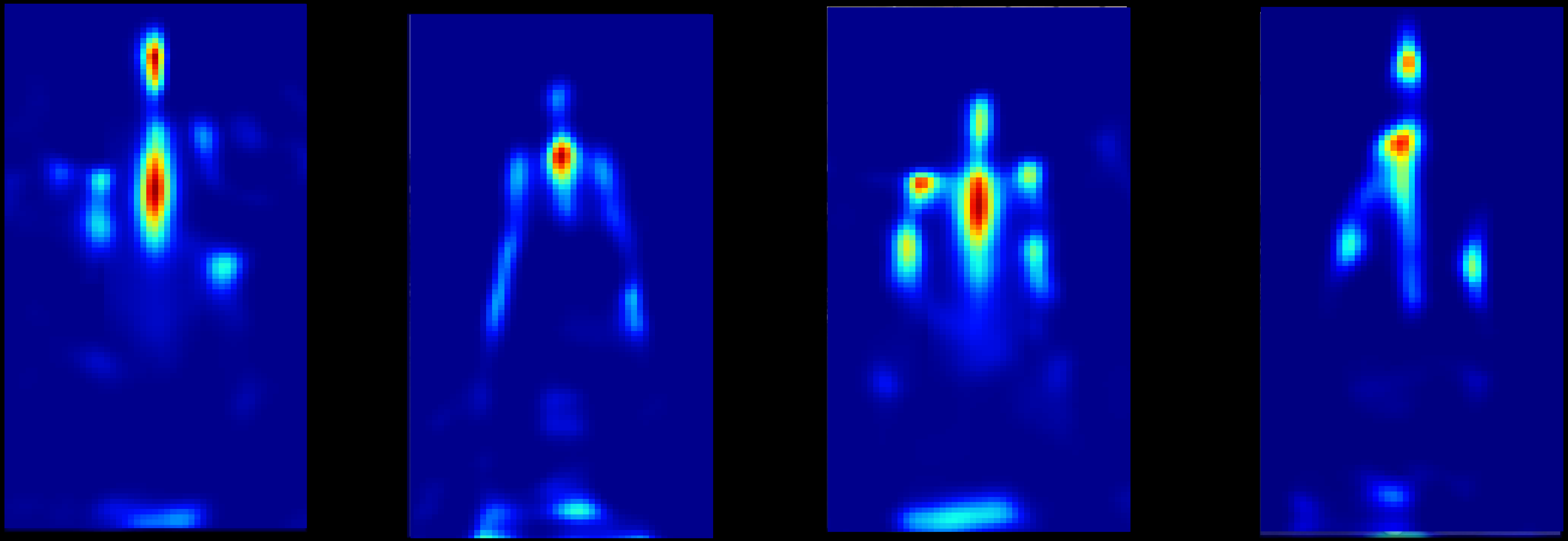
Evaluation



- RF-Capture sensor placed behind the wall
- 15 participants
- Use Kinect as baseline when needed

Sample Captured Figures through Walls

Sample Captured Figures through Walls

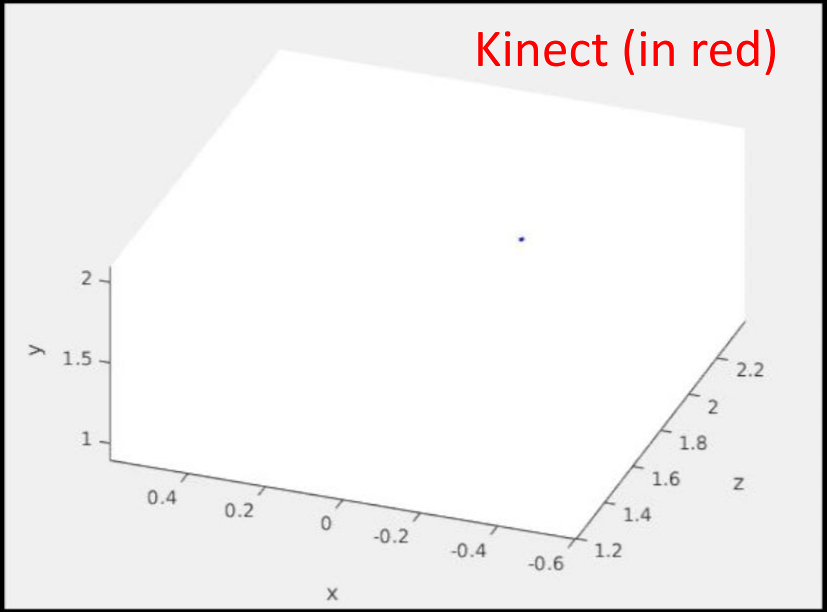


Tracking result

Writing in the air

Device

Our Tracking Result



Kinect (in red)

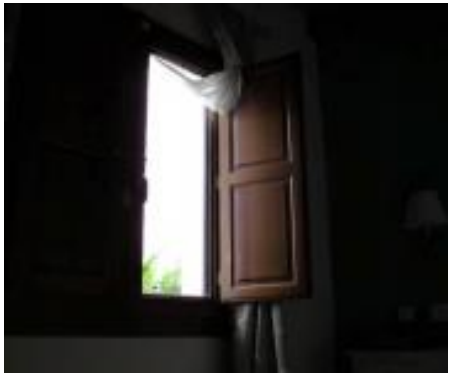
Median Accuracy is 2cm

Passive NLOS imaging using accidental pinholes

What does this image say about the world outside?

Accidental pinhole camera

Accidental pinhole camera



window is an aperture

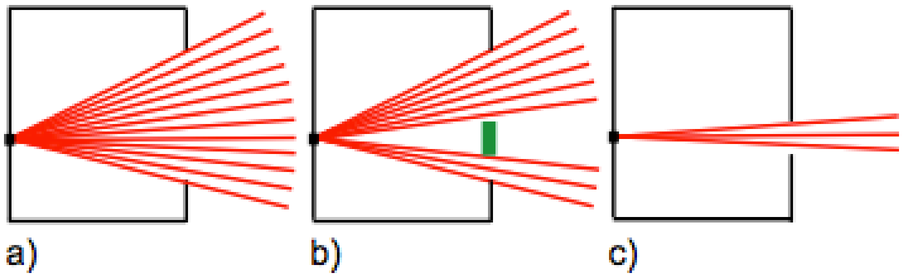
projected pattern on the wall

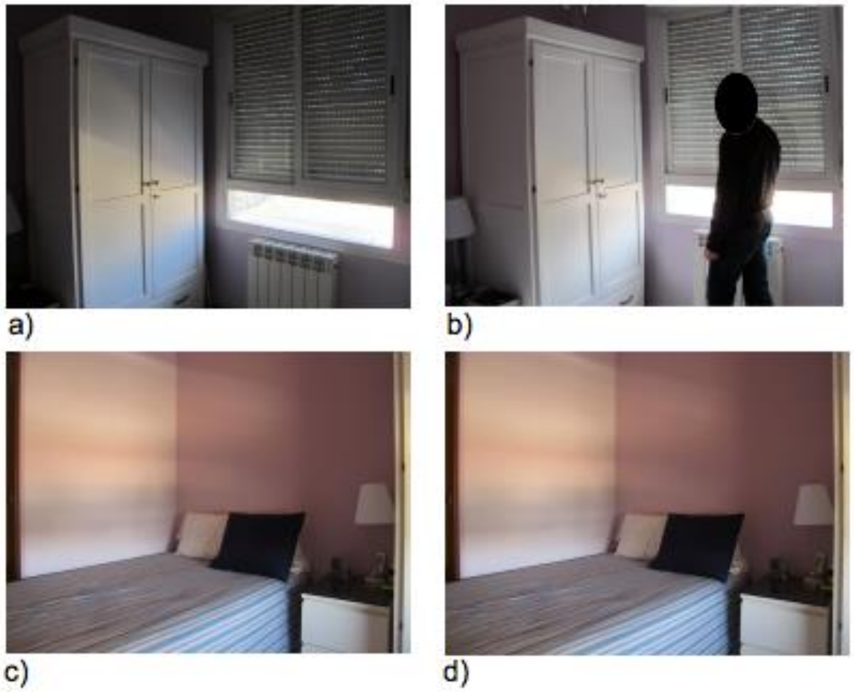
window with smaller gap

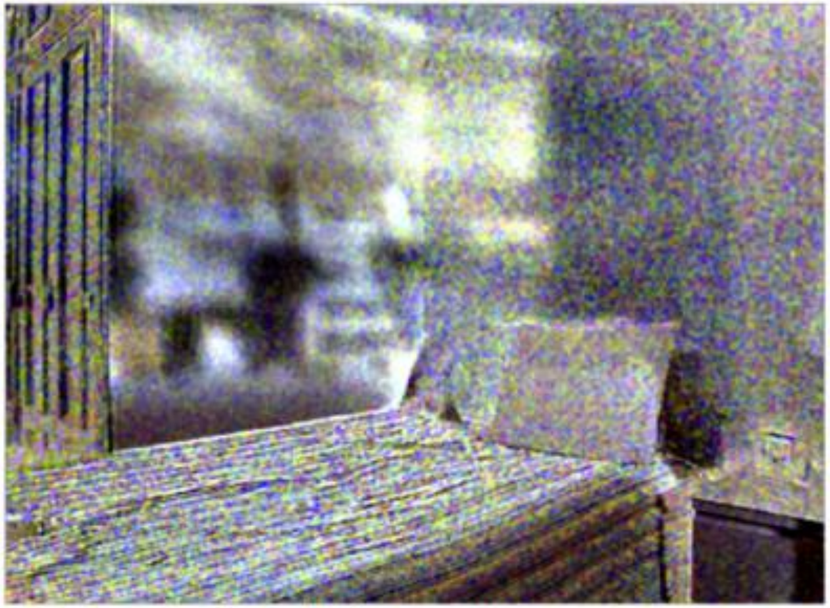
upside down

view outside window

Accidental pinspeck camera







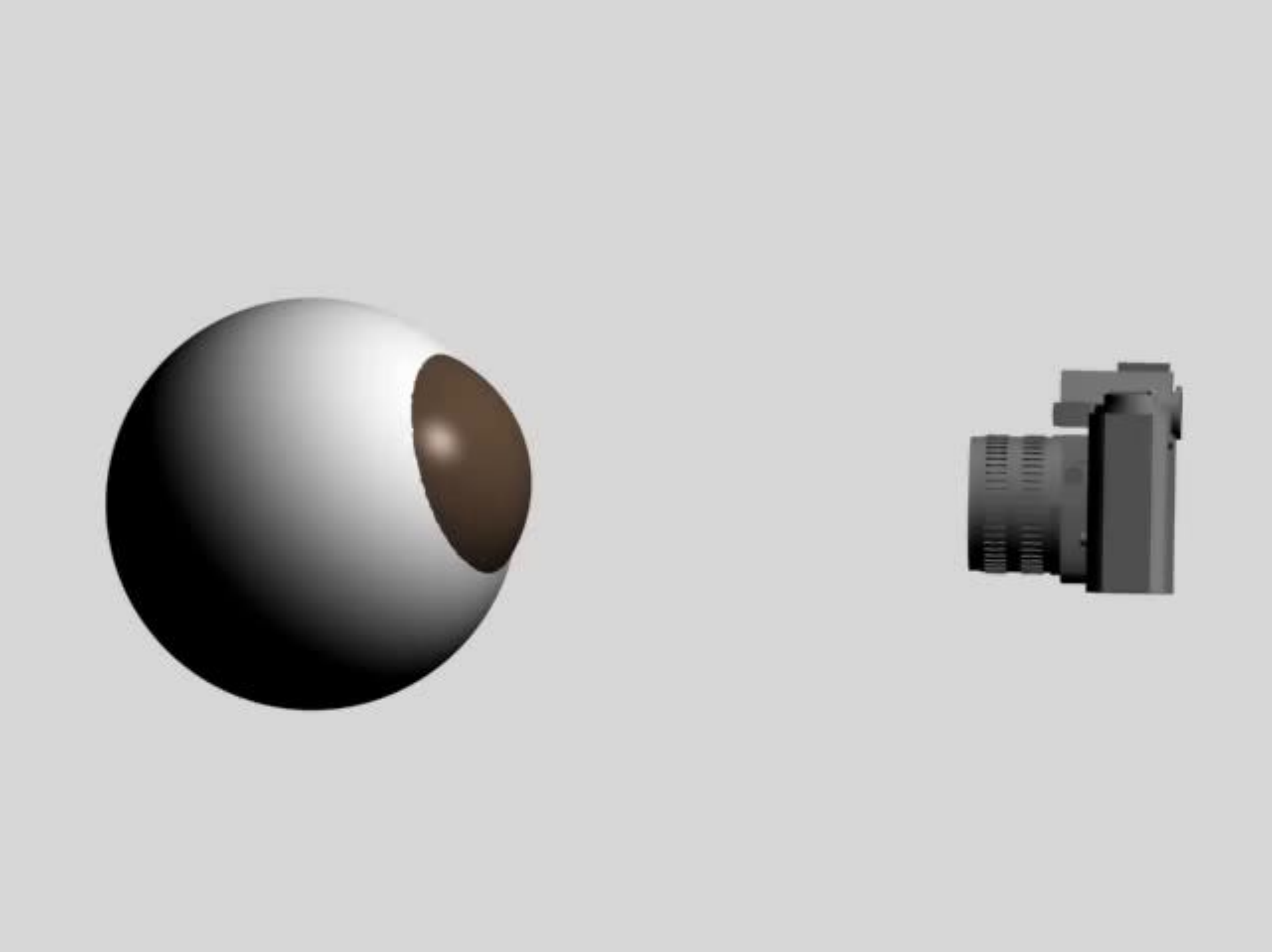
a) Difference image

b) Difference upside down

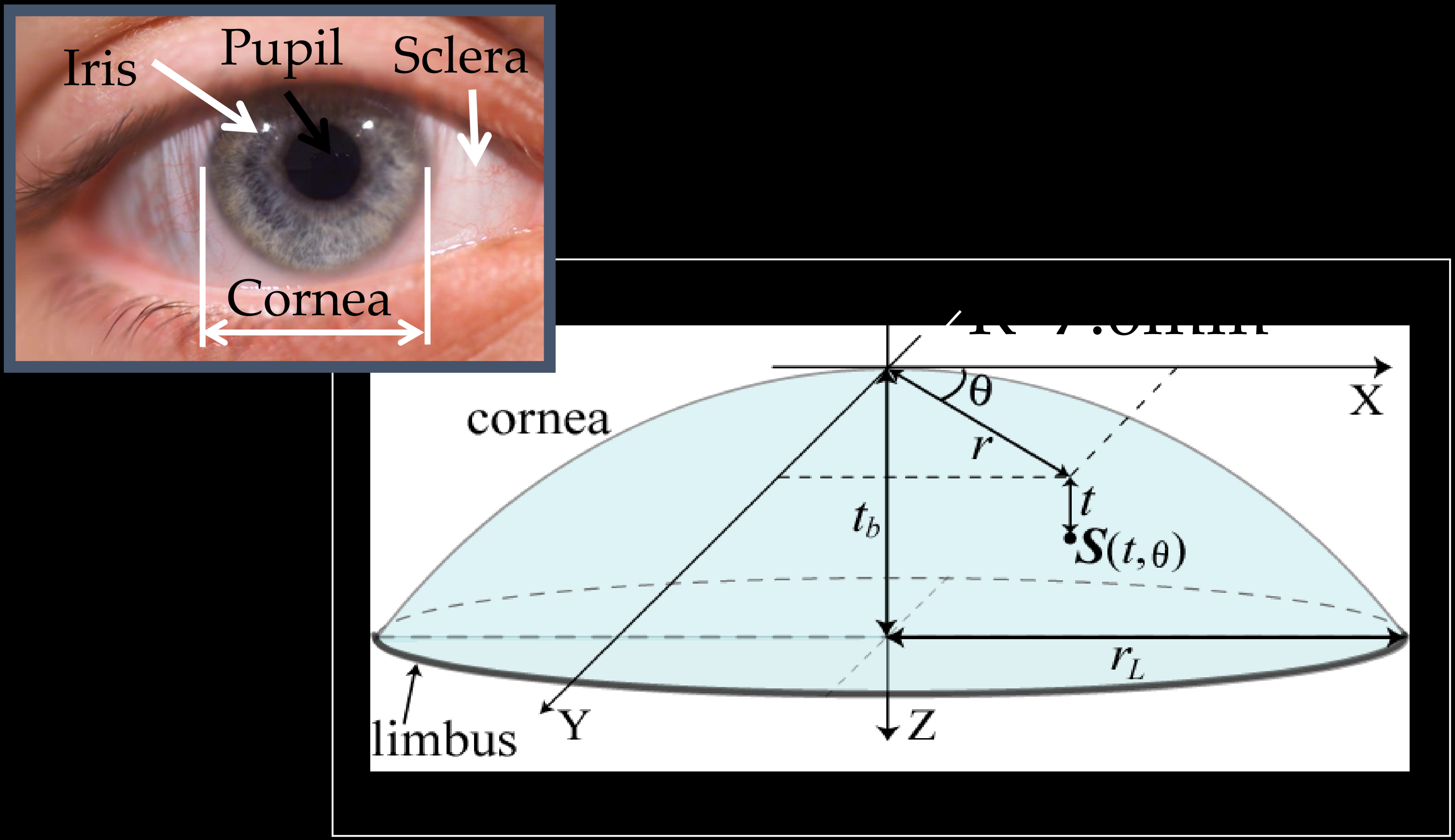
c) True outdoor view

Passive NLOS imaging using accidental reflectors

Corneal Imaging System



Geometric Model of the Cornea



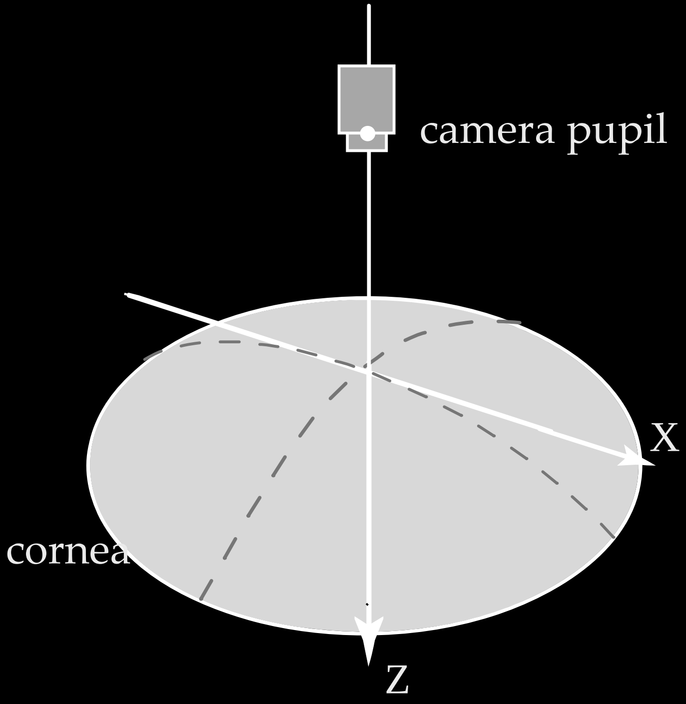
$$t_b = 2.18\text{mm} \qquad r_L = 5.5\text{mm}$$

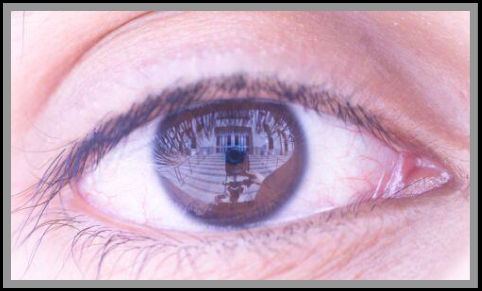
$$\text{eccentricity} = 0.5$$

Self-calibration:

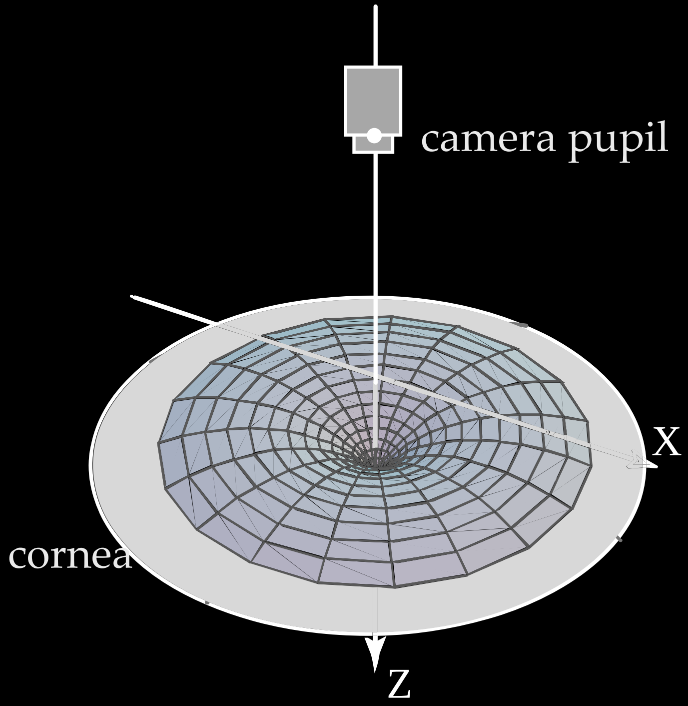
3D Coordinates, 3D Orientation

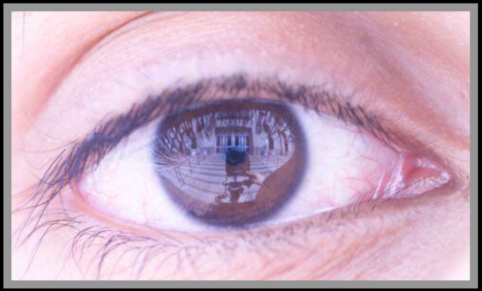
Viewpoint Loci



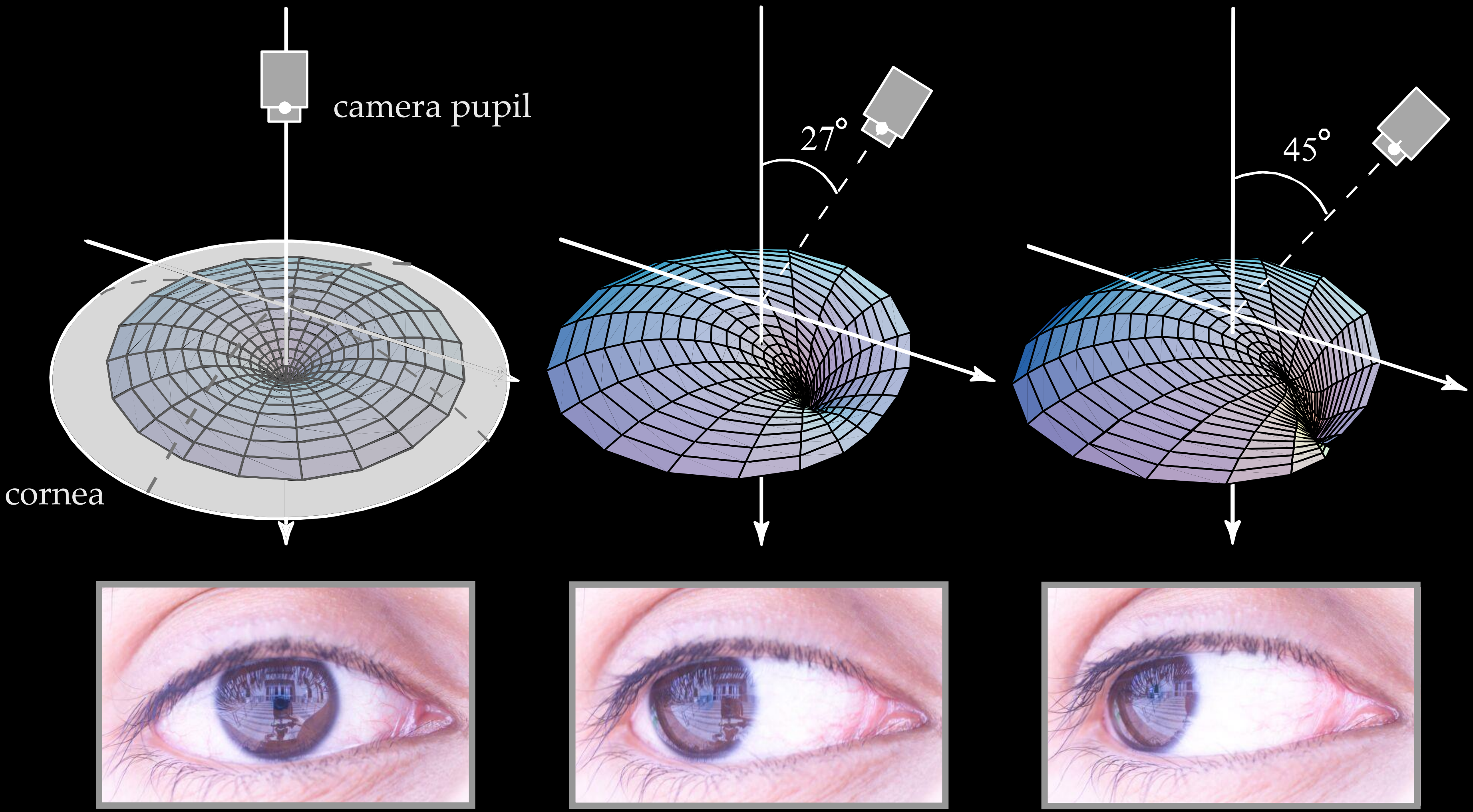


Viewpoint Loci

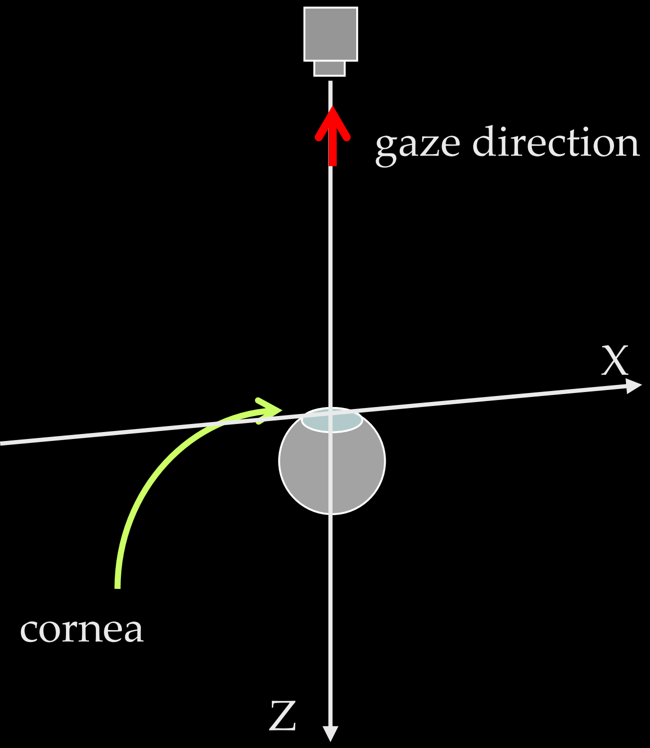


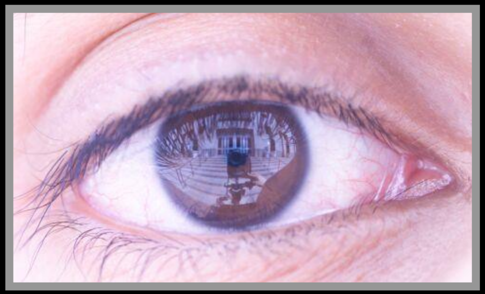


Viewpoint Loci

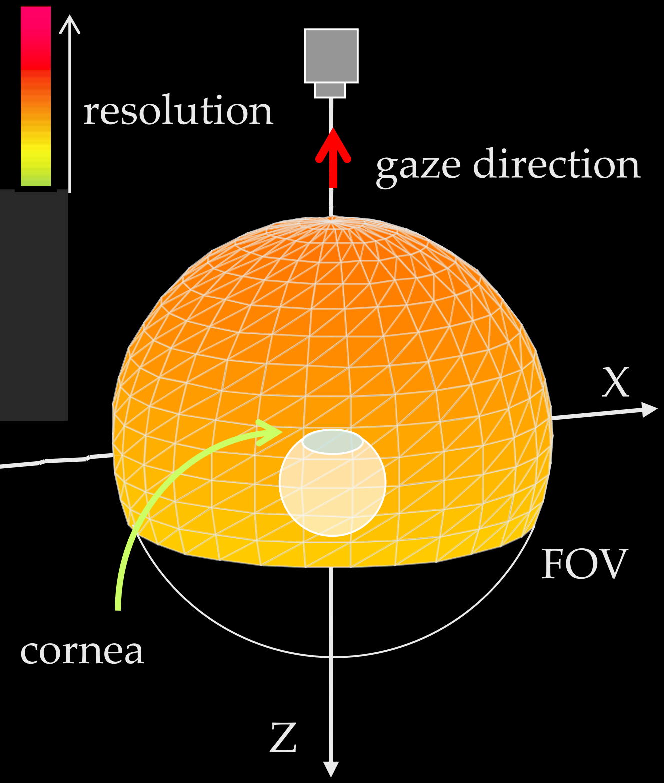


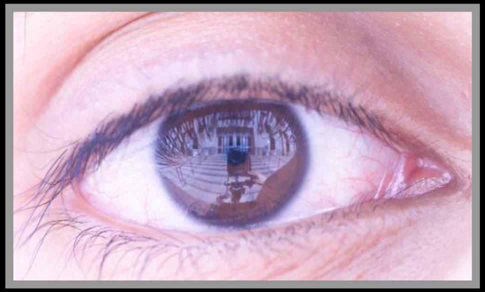
Resolution and Field of View



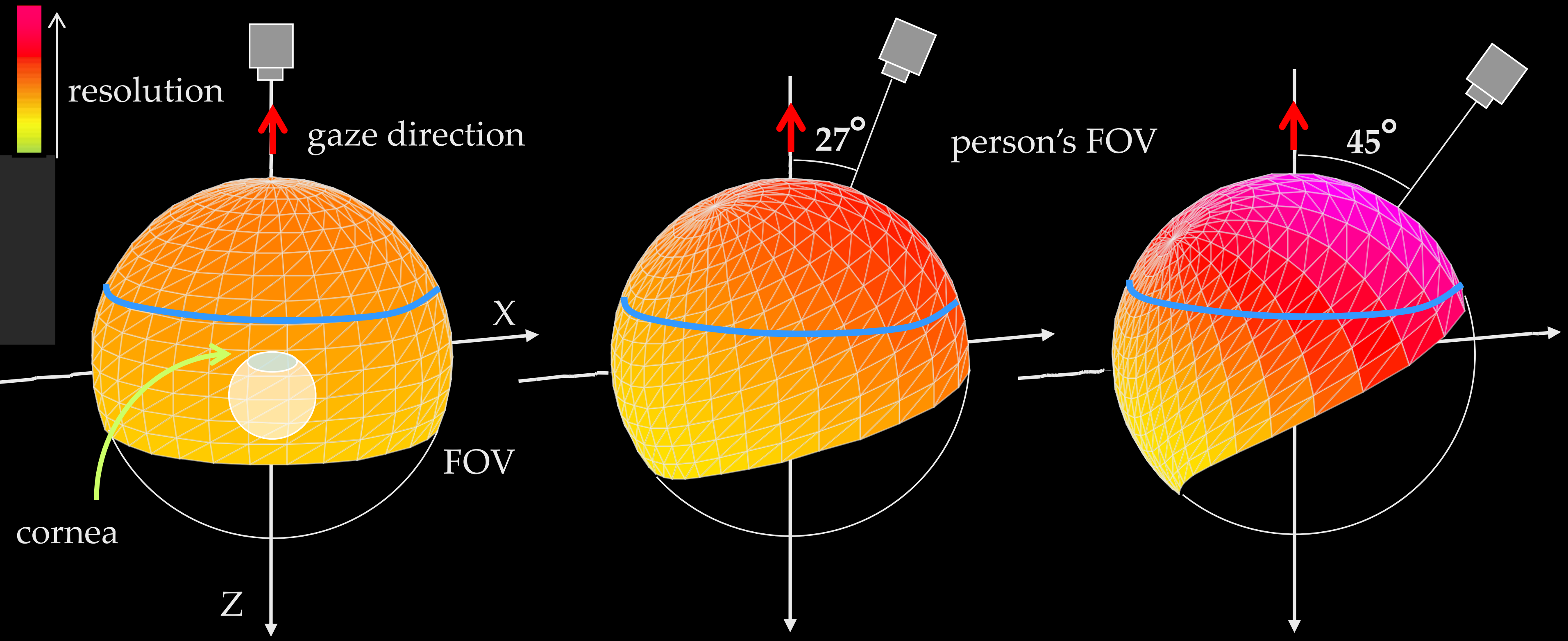


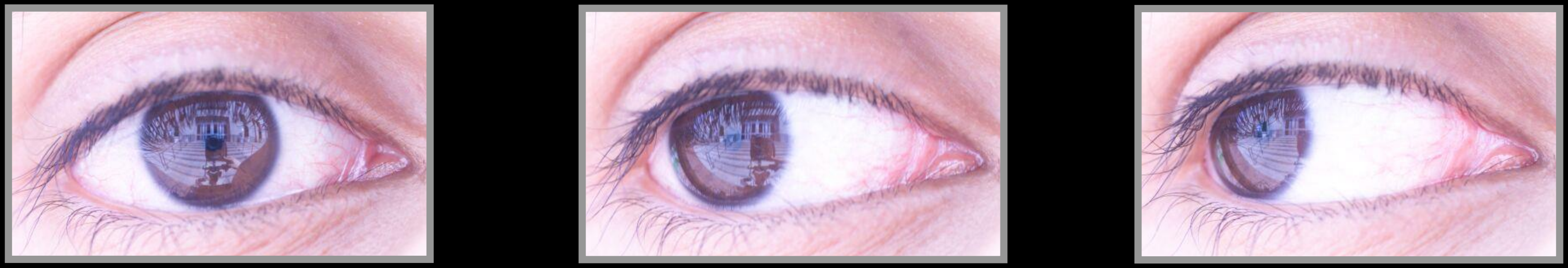
Resolution and Field of View

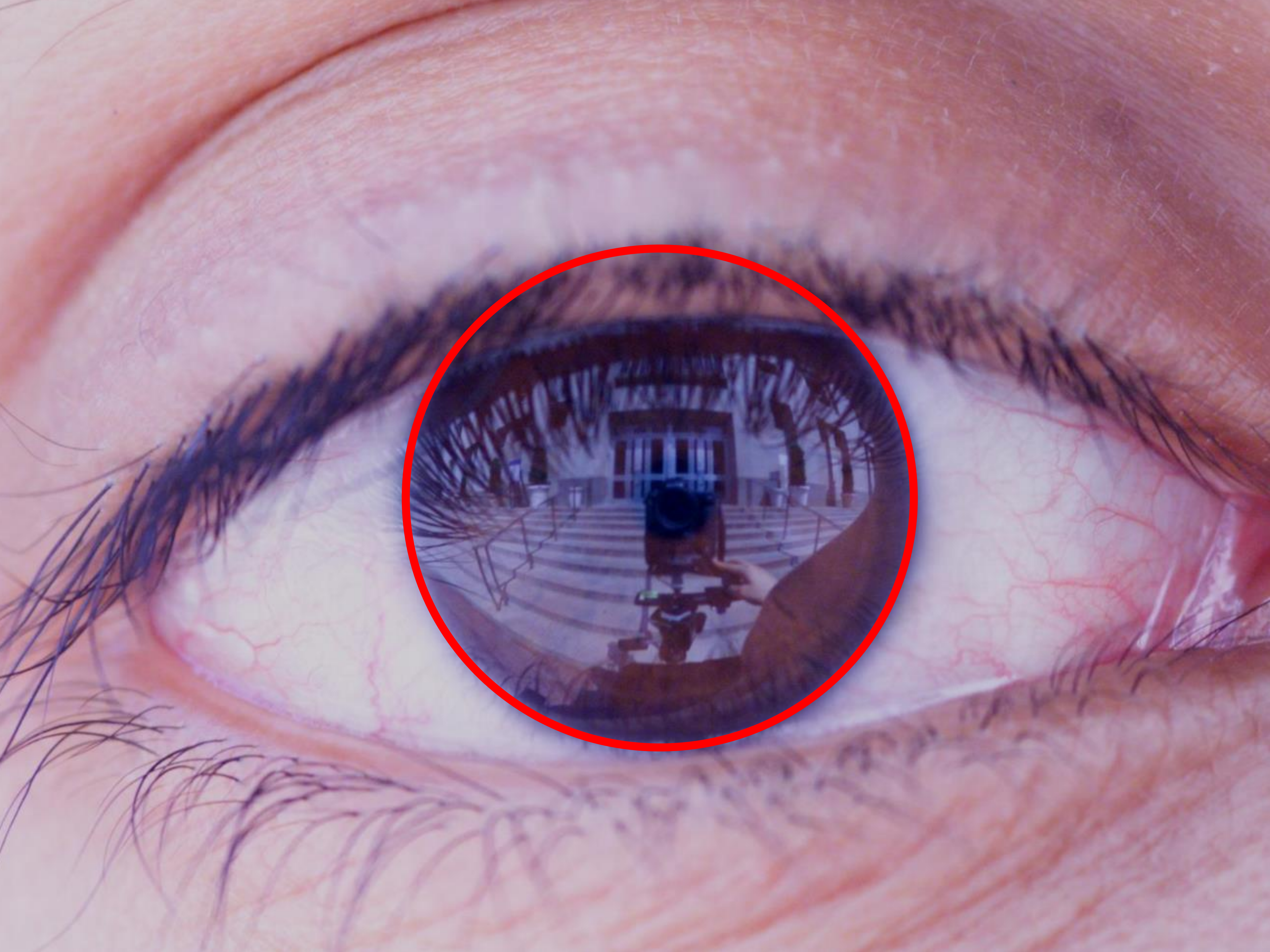




Resolution and Field of View





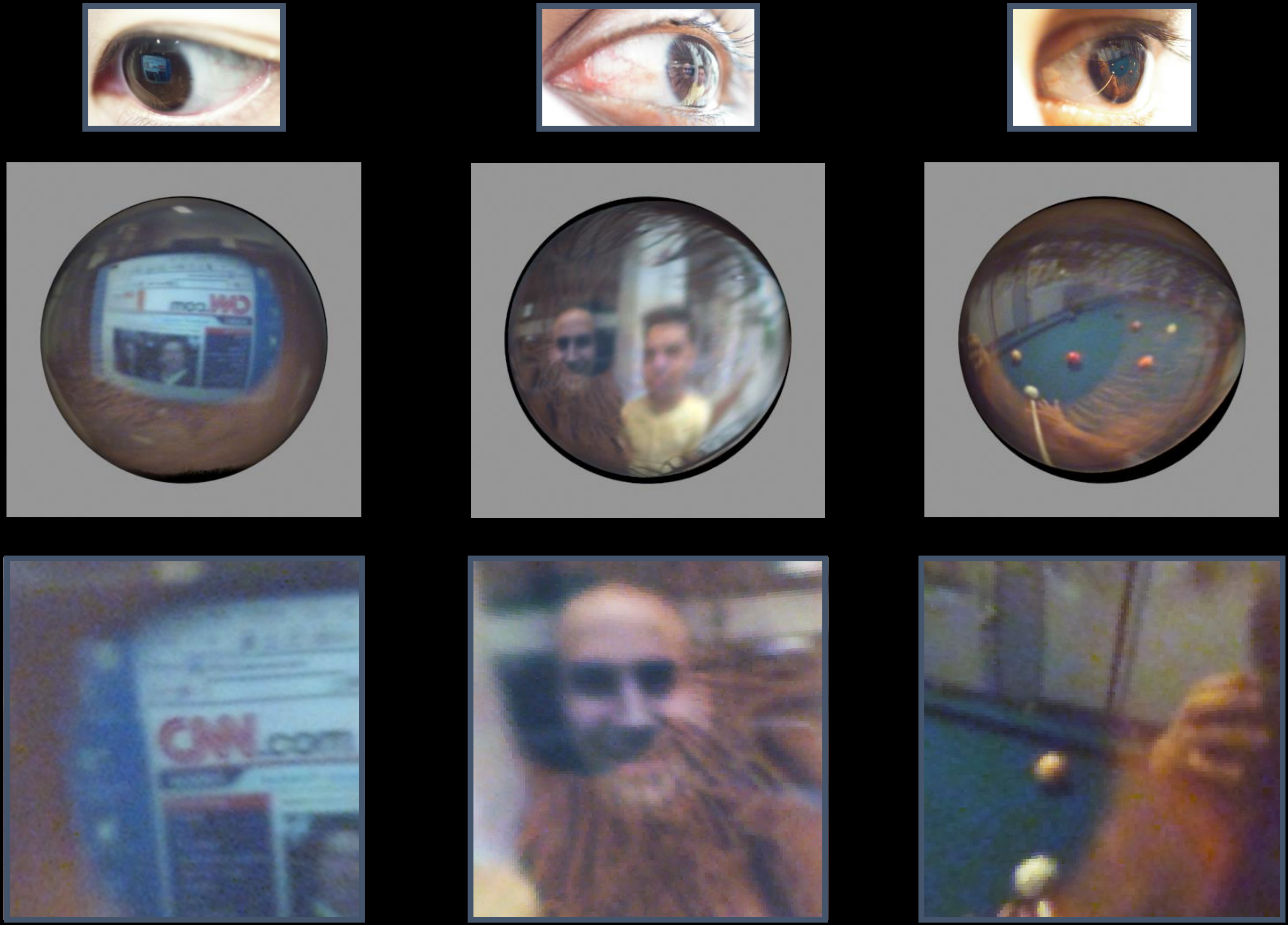


Environment Map from an Eye

What Exactly You are Looking At

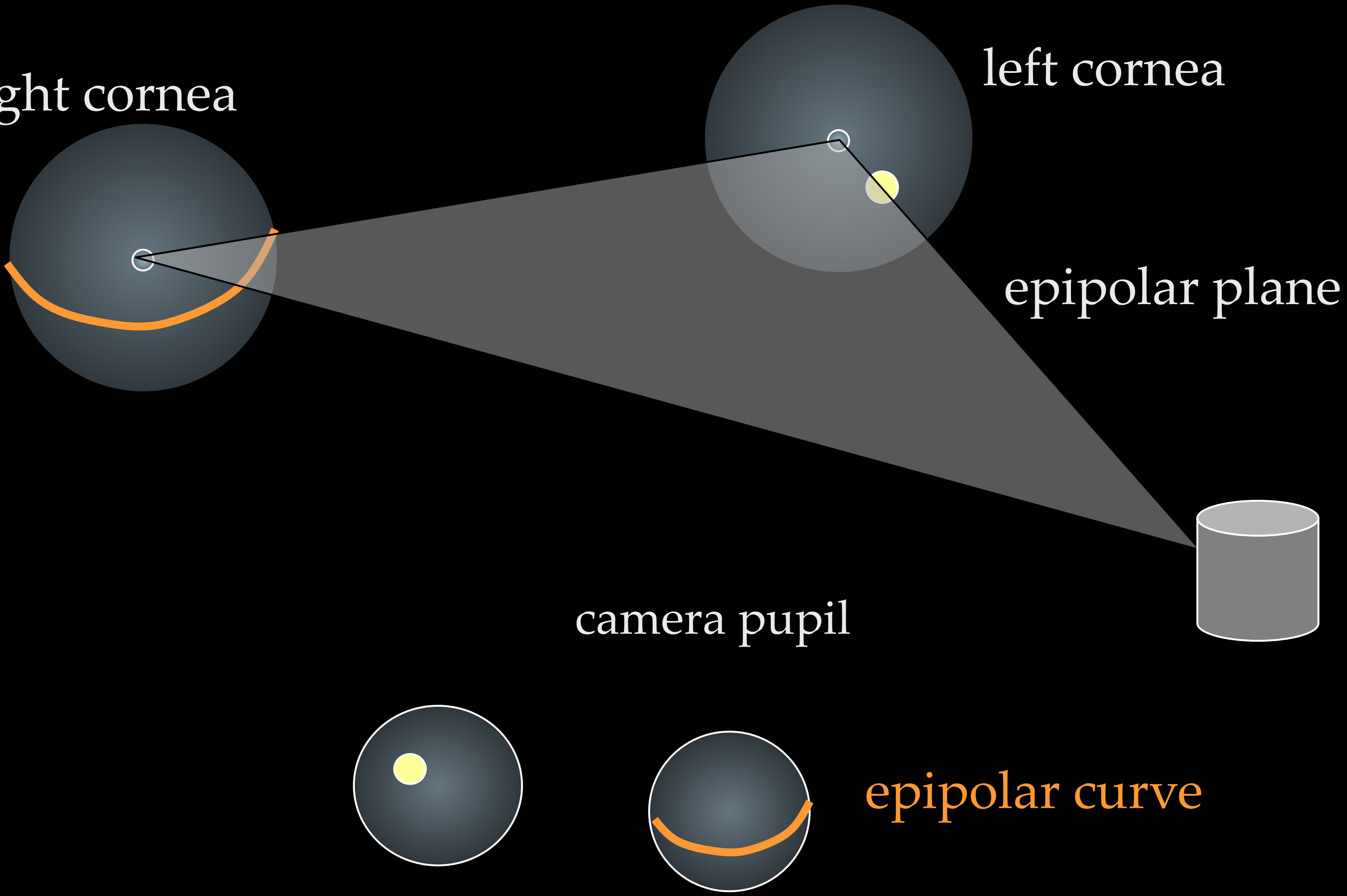
Eye Image:

Computed Retinal Image:

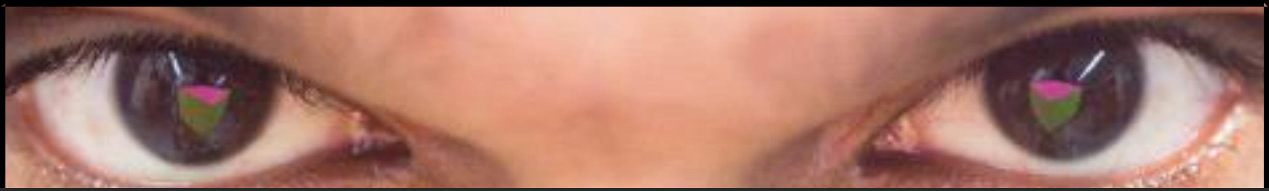

CNN.com
CNN.com

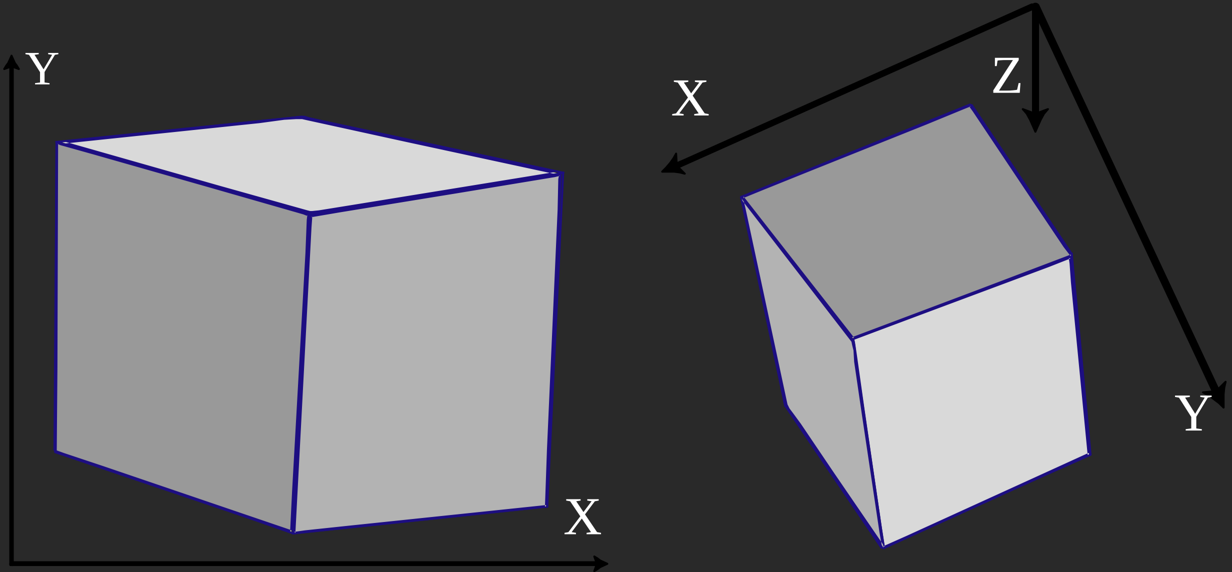
Watching a Bus

Corneal Stereo System



From Two Eyes in an Image …





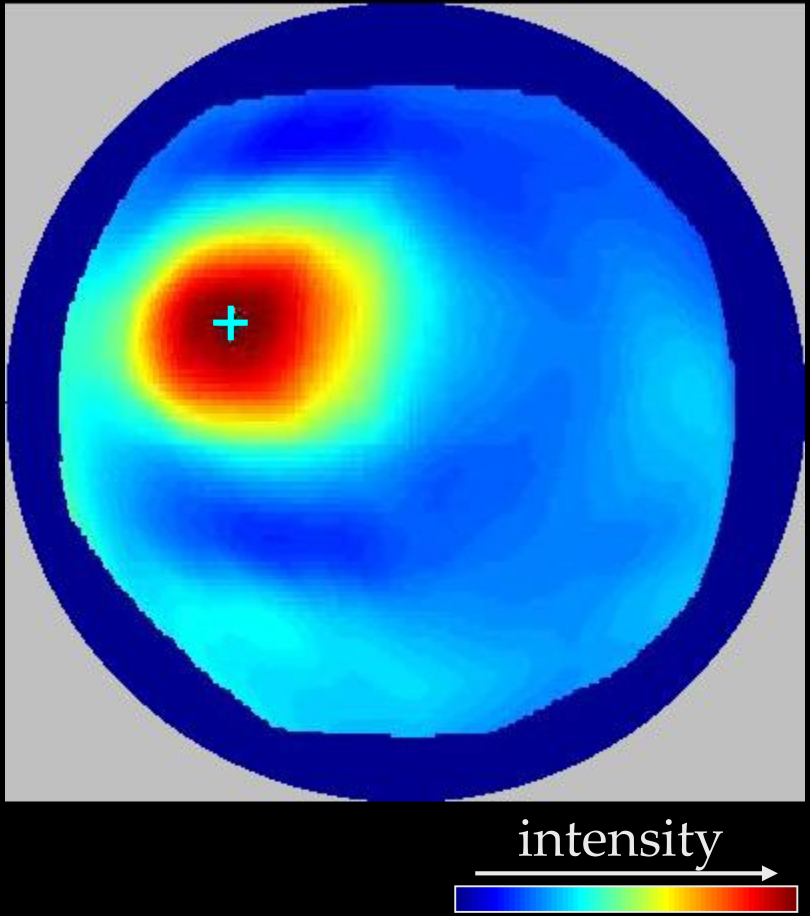
Reconstructed Structure (frontal and side view)

Eyes Reveal …

- Where the person is
- What the person is looking at
- The structure of objects

Dynamic Illumination in a Video

Point Source Direction from the Eye

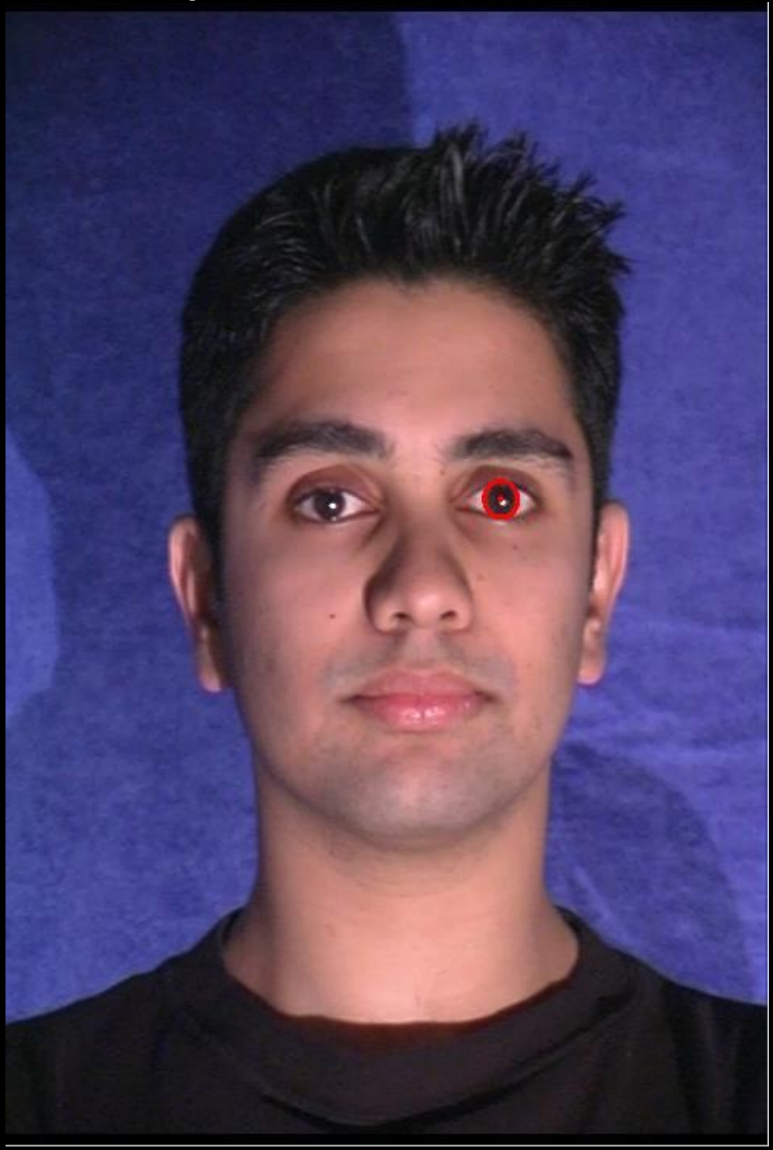


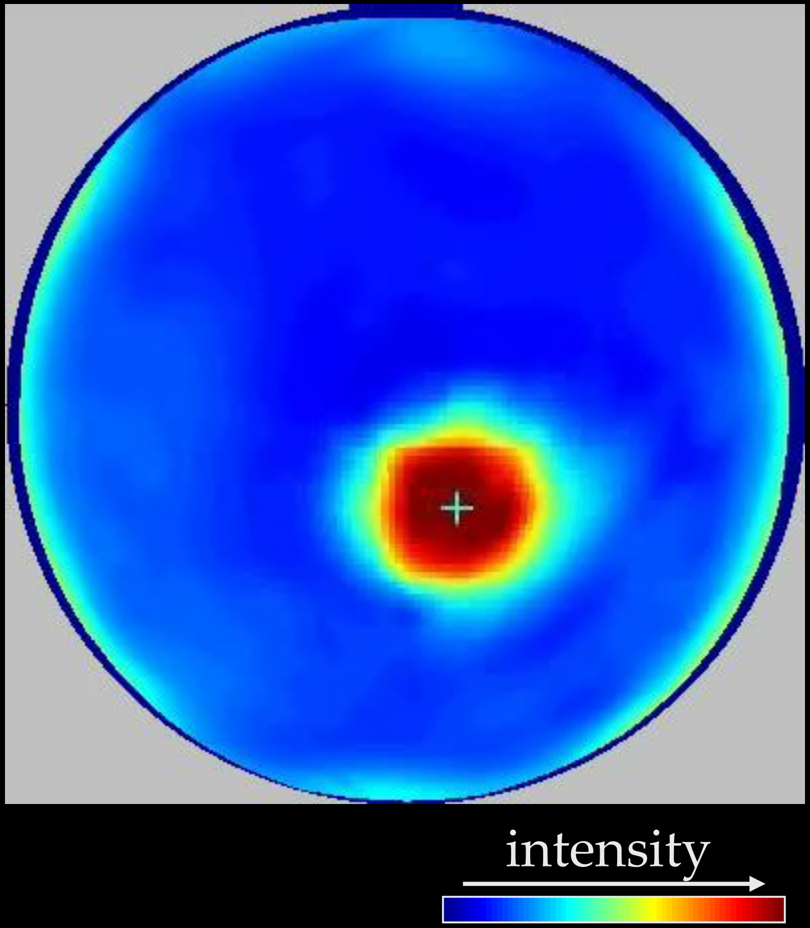
Inserting a Virtual Object

Sampling Appearance using Eyes

Sampling Appearance using Eyes

Computed Point Source Trajectory





Fitting a Reflectance Model

albedo map

normal map

3D Model Reconstruction

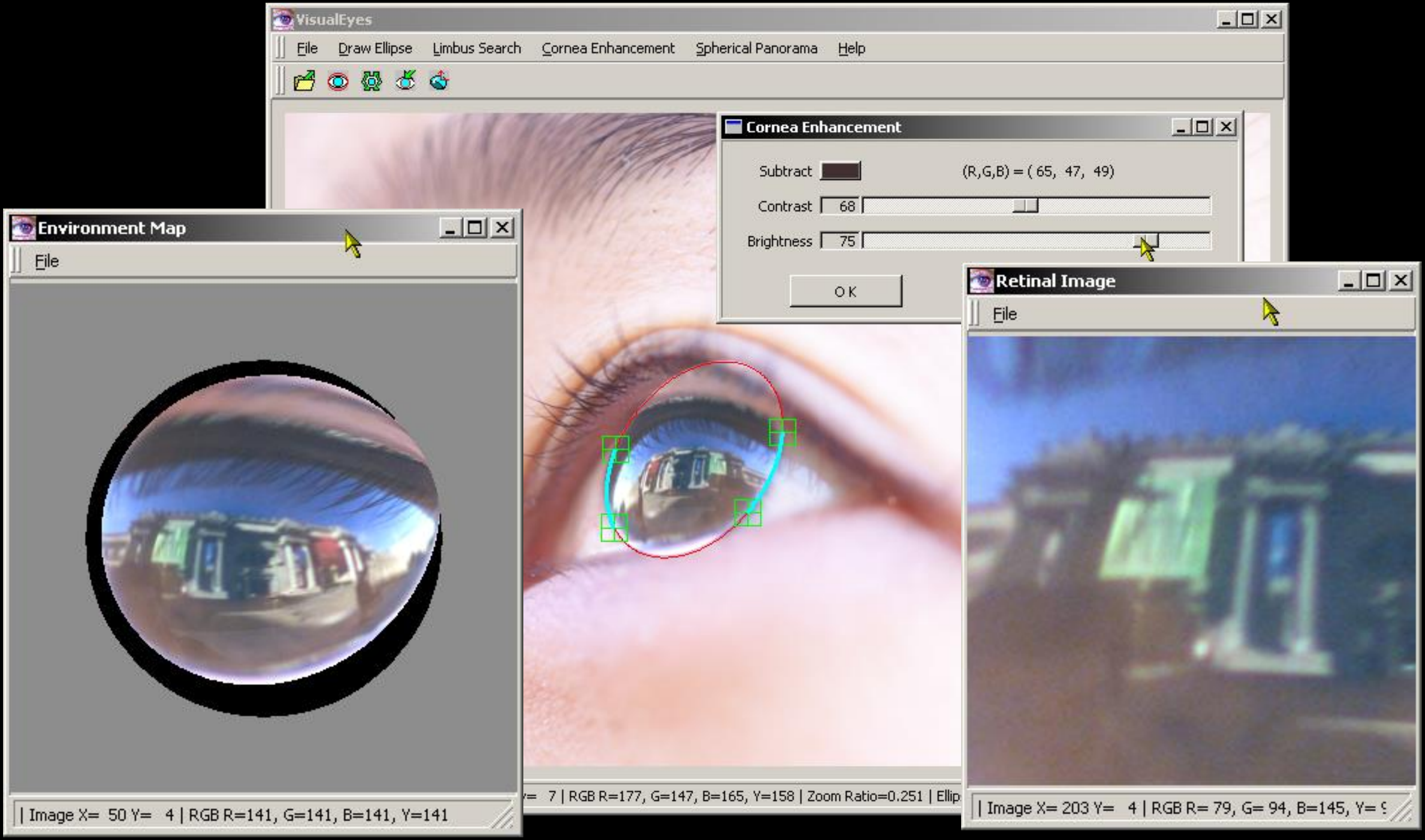
albedo map

3D model

Relighting under Novel Illumination

VisualEyes™

http://www.cs.columbia.edu/CAVE/



Cornea Enhancement

Subtract (R,G,B) = (65, 47, 49)

Contrast 68

Brightness 75

OK

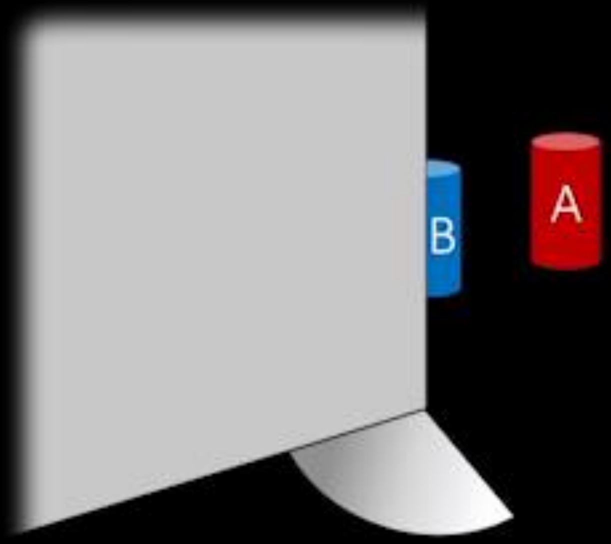
```
| Image X=  50 Y=   4 | RGB R=141, G=141, B=141, Y=141
=   7 | RGB R=177, G=147, B=165, Y=158 | Zoom Ratio=0.251 | Ellip
| Image X= 203 Y=   4 | RGB R= 79, G= 94, B=145, Y=
```

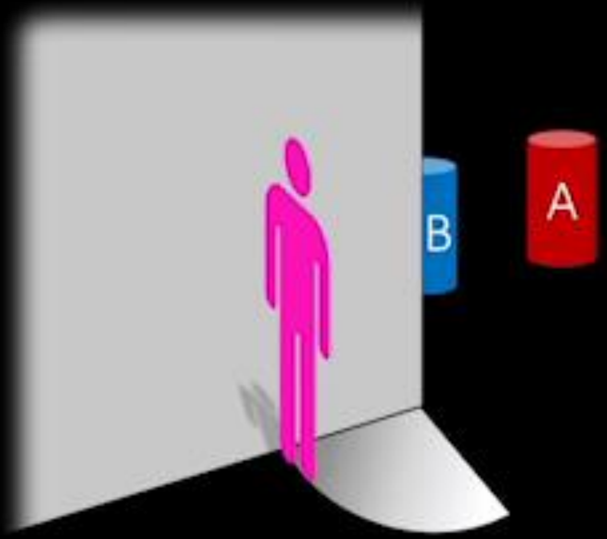
with Akira Yanagawa

Passive NLOS imaging using corners

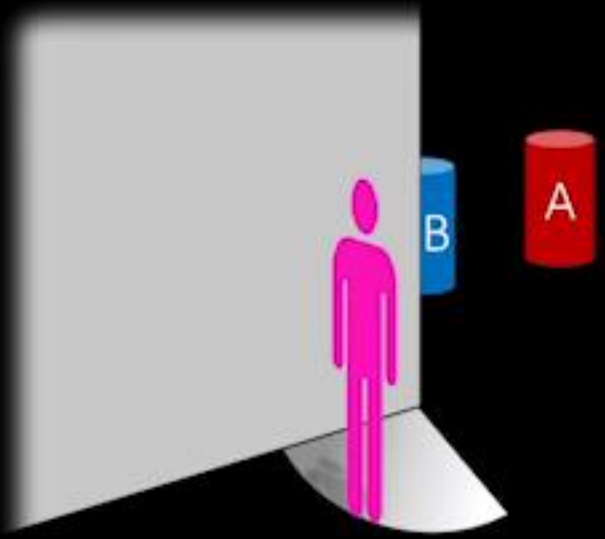
Hidden Scene

Video of the Corner

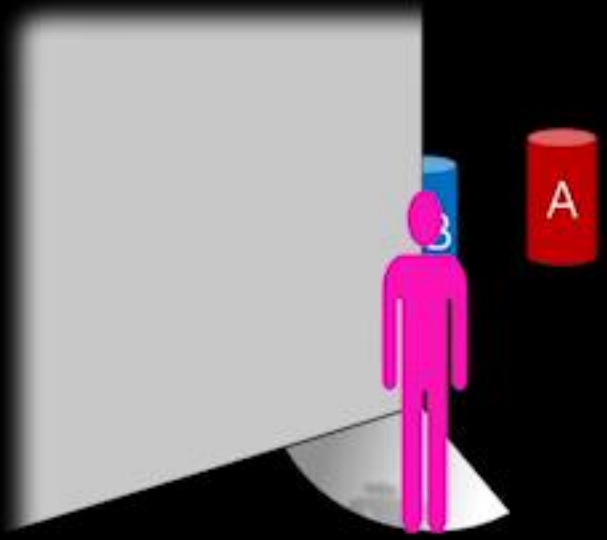

B
A



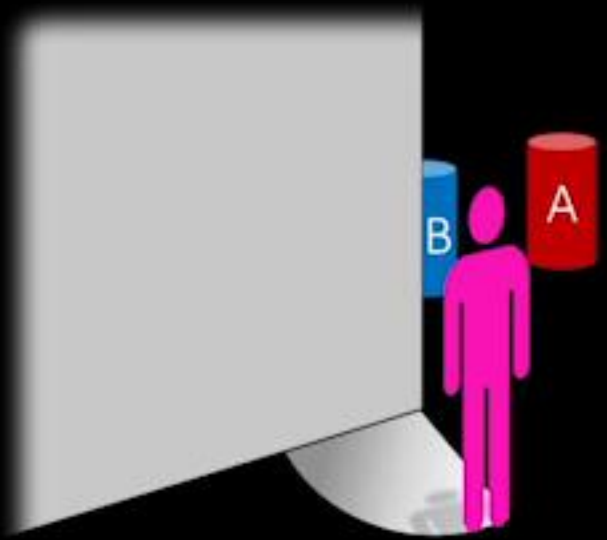
What You Would See



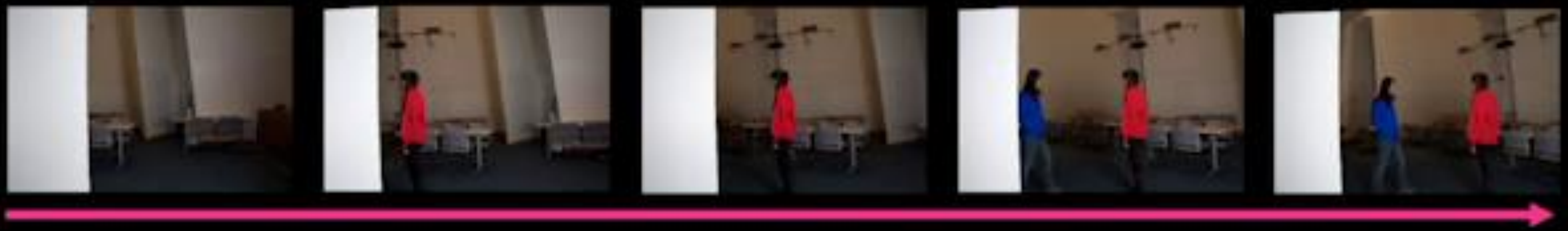
What You Would See

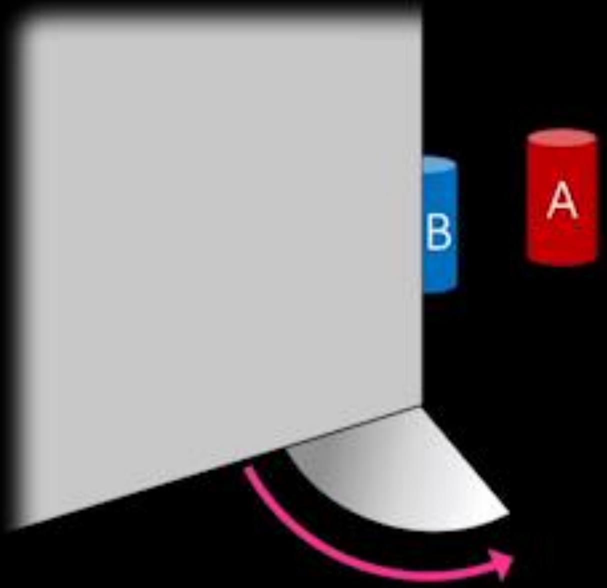


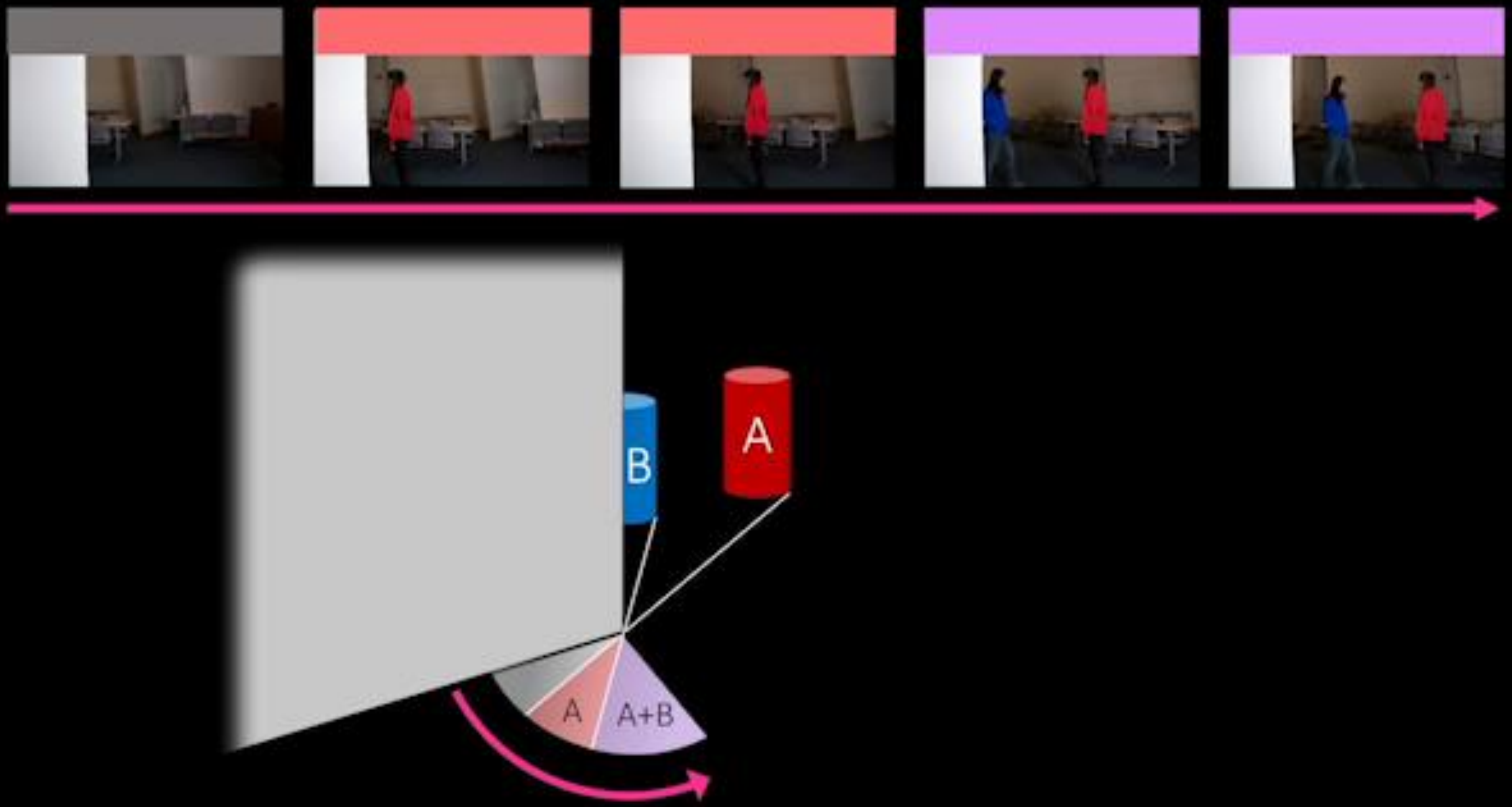

What You Would See

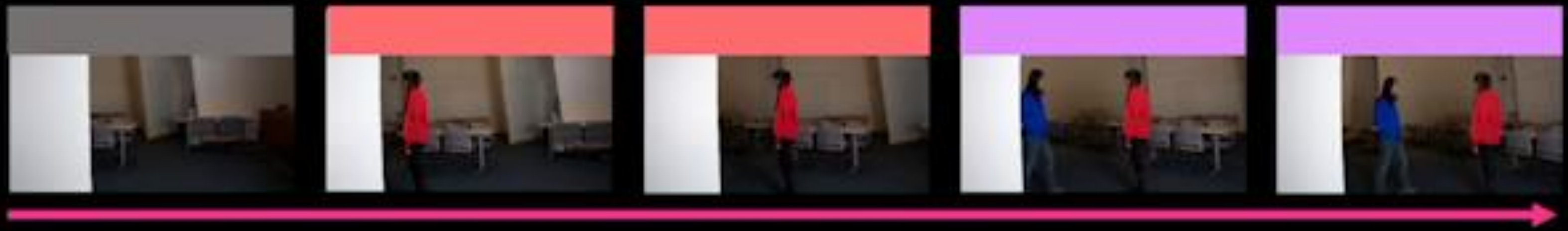


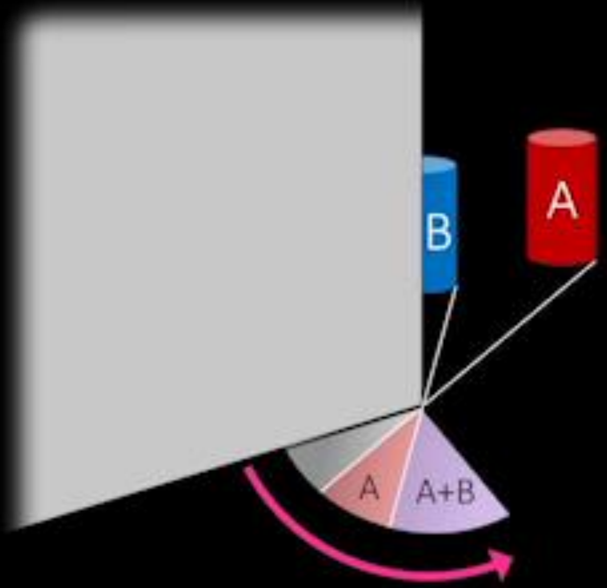
What You Would See



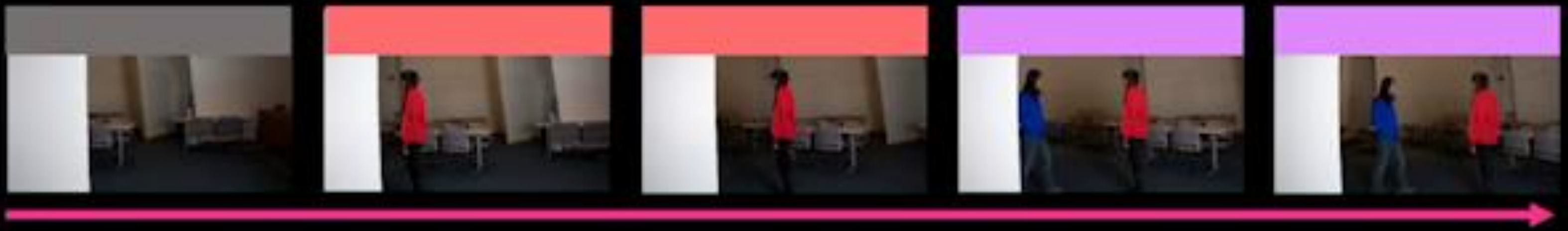

B
A

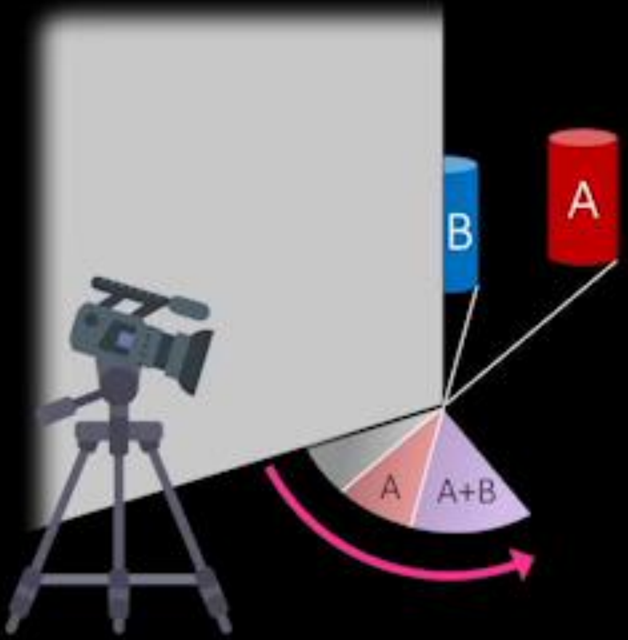

B
A
A
A+B



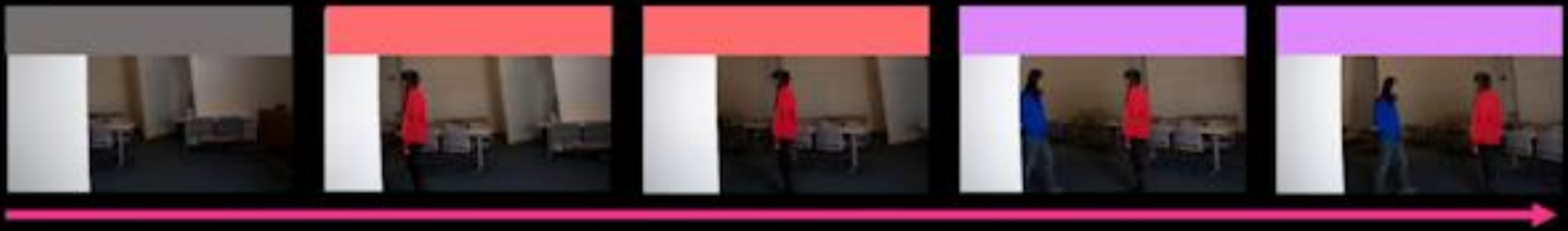


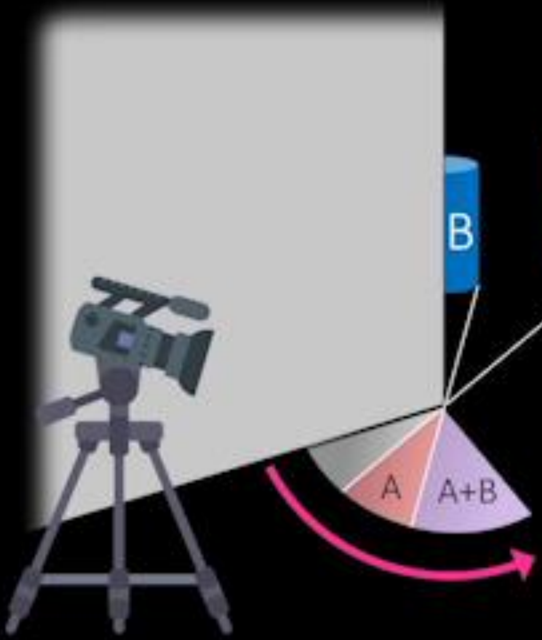
A person makes a 0.1% difference in the reflected light at the base of a corner





Original Frame

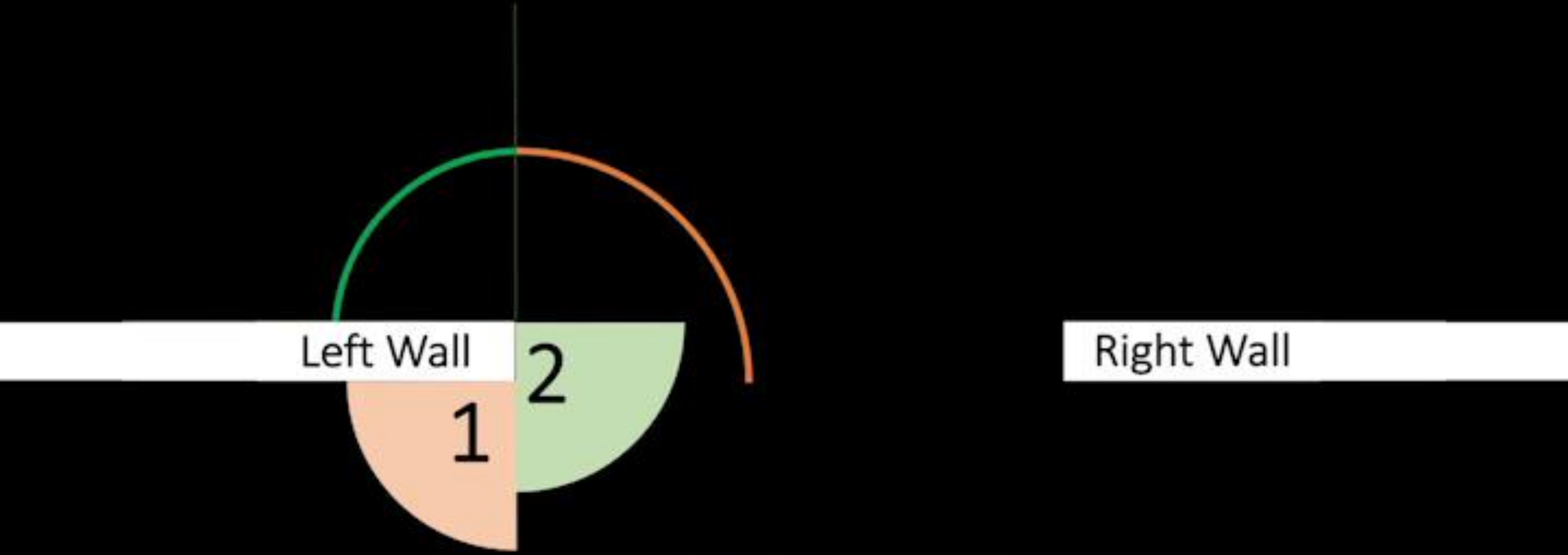




Color Magnified

Left Wall

Right Wall

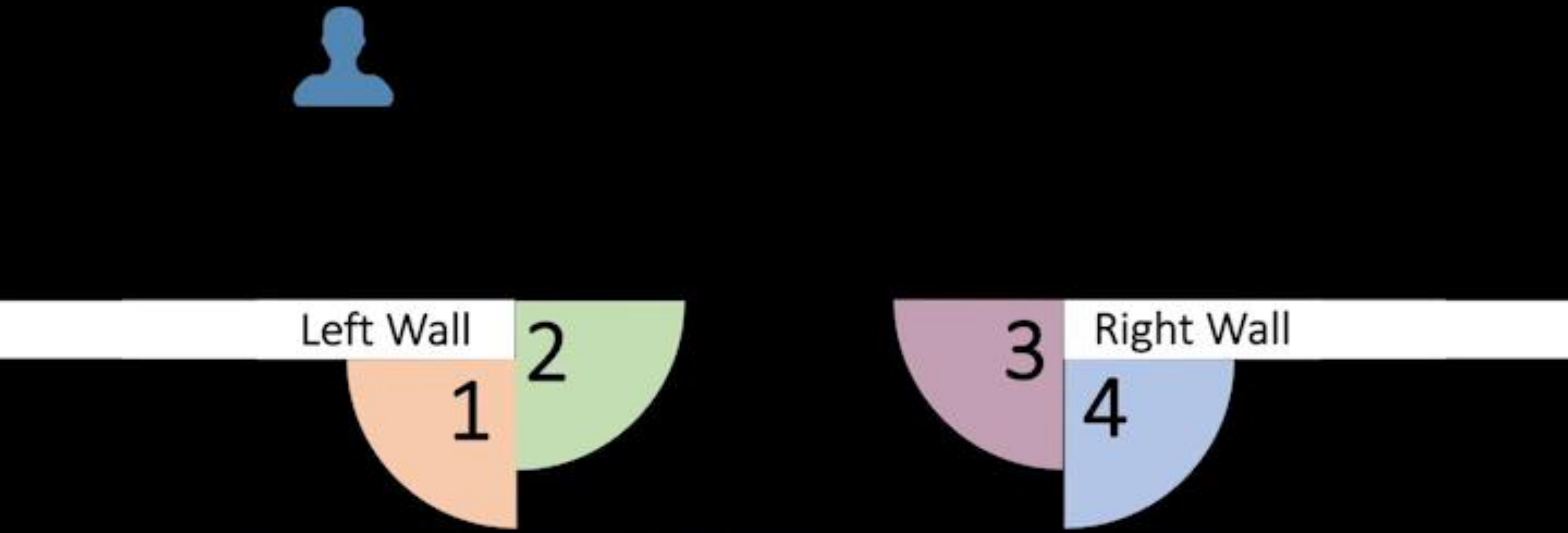


Left Wall

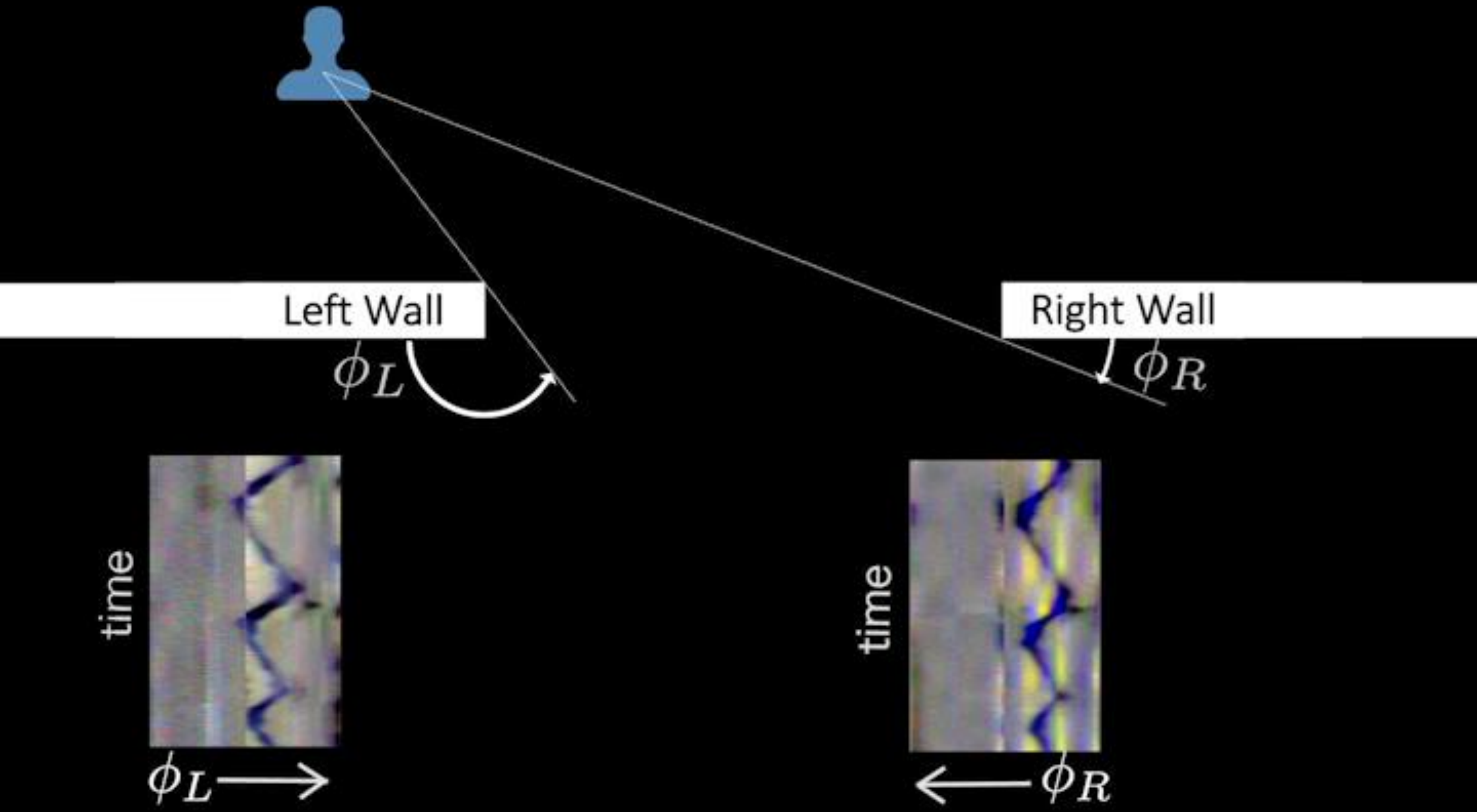
Right Wall

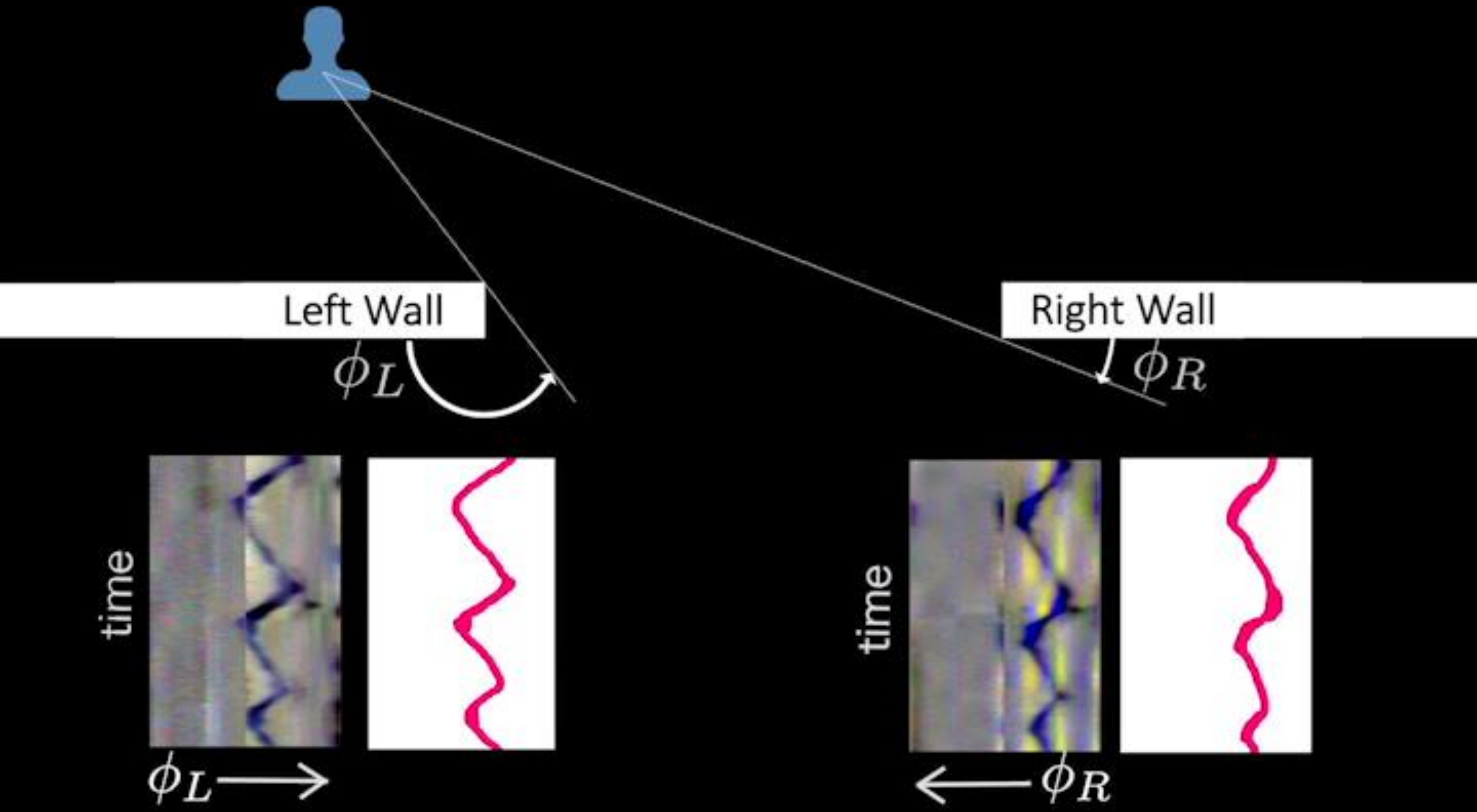

Left Wall
2
1
3
Right Wall
4

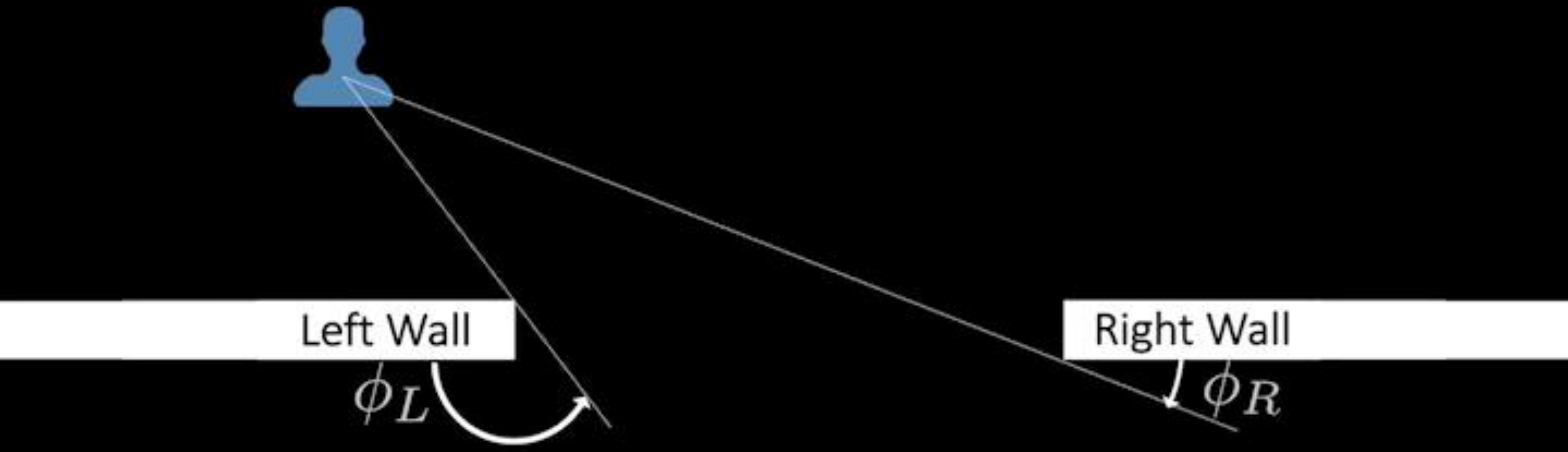

2
1
3
4


Left Wall
2
1
3
Right Wall
4


2
1
3
4


Left Wall
Right Wall
ϕ_L
ϕ_R
time
time
$\phi_L \longrightarrow$
$\longleftarrow \phi_R$


Left Wall
Right Wall
ϕ_L
ϕ_R
time
time
$\phi_L \longrightarrow$
$\longleftarrow \phi_R$




Full Scene


Zoom-in on Stereo Edge Cameras

References

Basic reading:

- Kirmani et al., “Looking around the corner using ultrafast transient imaging,” ICCV 2009 and IJCV 2011.
- Velten et al., “Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging,” Nature Communications 2012.
- O’Toole et al., “Confocal non-line-of-sight imaging,” Nature 2018.

  The first two papers introduced the idea of using ToF imaging for looking around the corner. The third paper discusses the advantages of the confocal scanning case.
- Abib and Katabi, “See Through Walls with Wi-Fi!,” SIGCOMM 2013.
- Abib et al., “Capturing the Human Figure Through a Wall,” SIGGRAPH Asia 2015.

  The two papers showing that WiFi can be used to see through walls.
- Torralba and Freeman, “Accidental Pinhole and Pinspeck Cameras,” CVPR 2012.

  The paper discussing passive NLOS imaging using accidental pinholes.
- Nishimo and Nayar, “Corneal Imaging System: Environment from Eyes,” IJCV 2006.

  The paper discussing passive NLOS imaging using accidental reflectors.
- Bouman et al., “Turning corners into cameras: Principles and Methods,” ICCV 2017.

  The paper discussing passive NLOS imaging using corners.

Additional reading:

- Pediredla et al., “Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner,” ICCP 2017.

  The paper on NLOS room reconstruction using ToF imaging.
- Nishimo and Nayar, “Eyes for relighting,” SIGGRAPH 2004.

  A follow-up paper to the paper on corneal imaging, show how similar ideas can be used for relighting and other image-based rendering tasks.