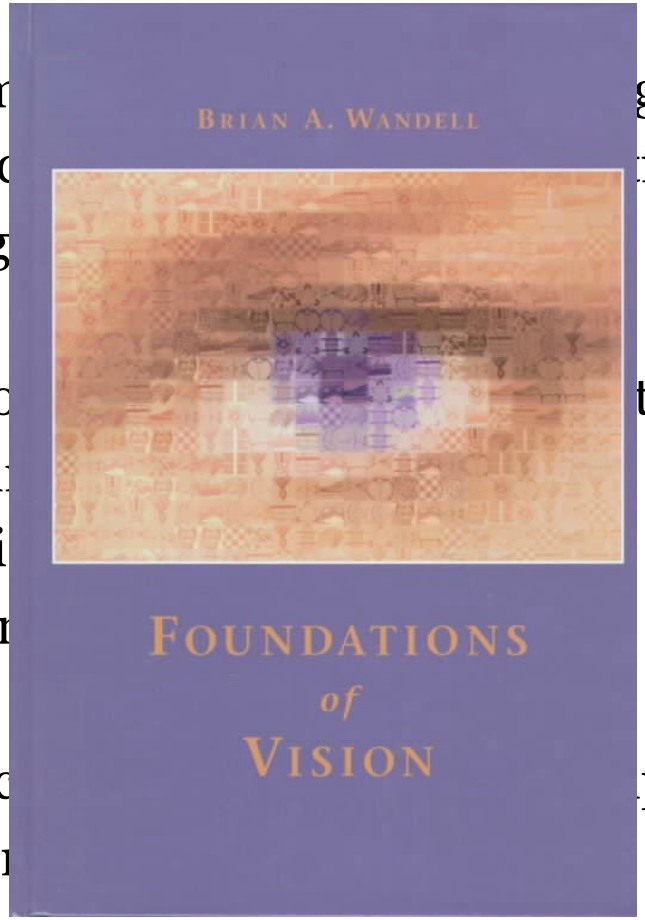
Simulation technologies for image systems engineering

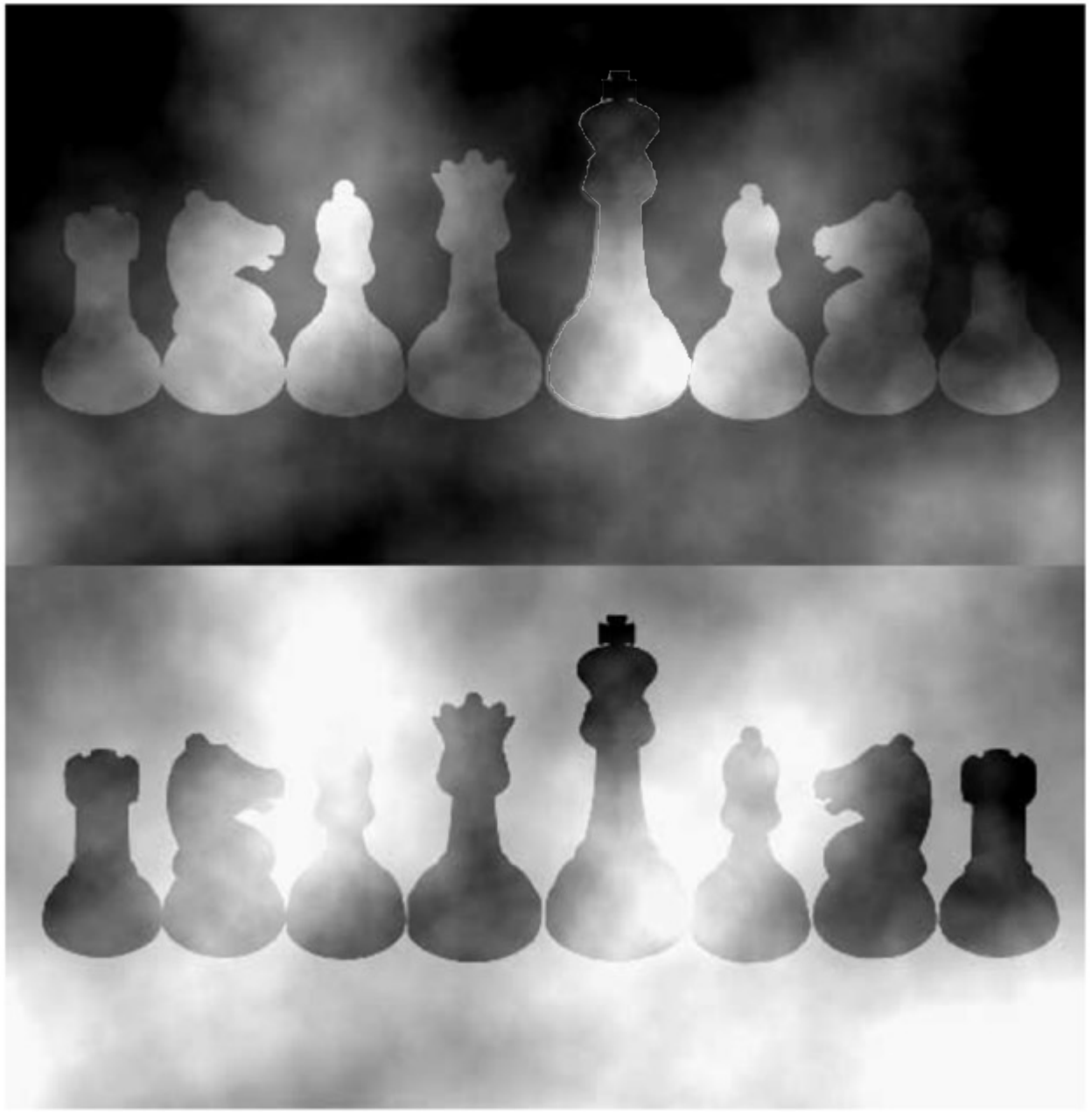
Brian A. Wandell
Stanford Center for Image Systems Engineering (SCIEN)

Wu Tsai Neurosciences Institute (WTNI)
Stanford Center for Cognitive and Neurobiological Imaging (CNI)

Vision science

- Even sim ghtness - depend ion of the imag circuits
- The visio tial in developi neurosci intelliger
- Vision sc portant for the e





Anderson and Winawer (Nature, 2005)

Studying the brain (Psych 204a)

- Brain computations depend on a variety of cells; one important cell type, the neurons, have their cell bodies located in the cerebral cortex (gray matter).

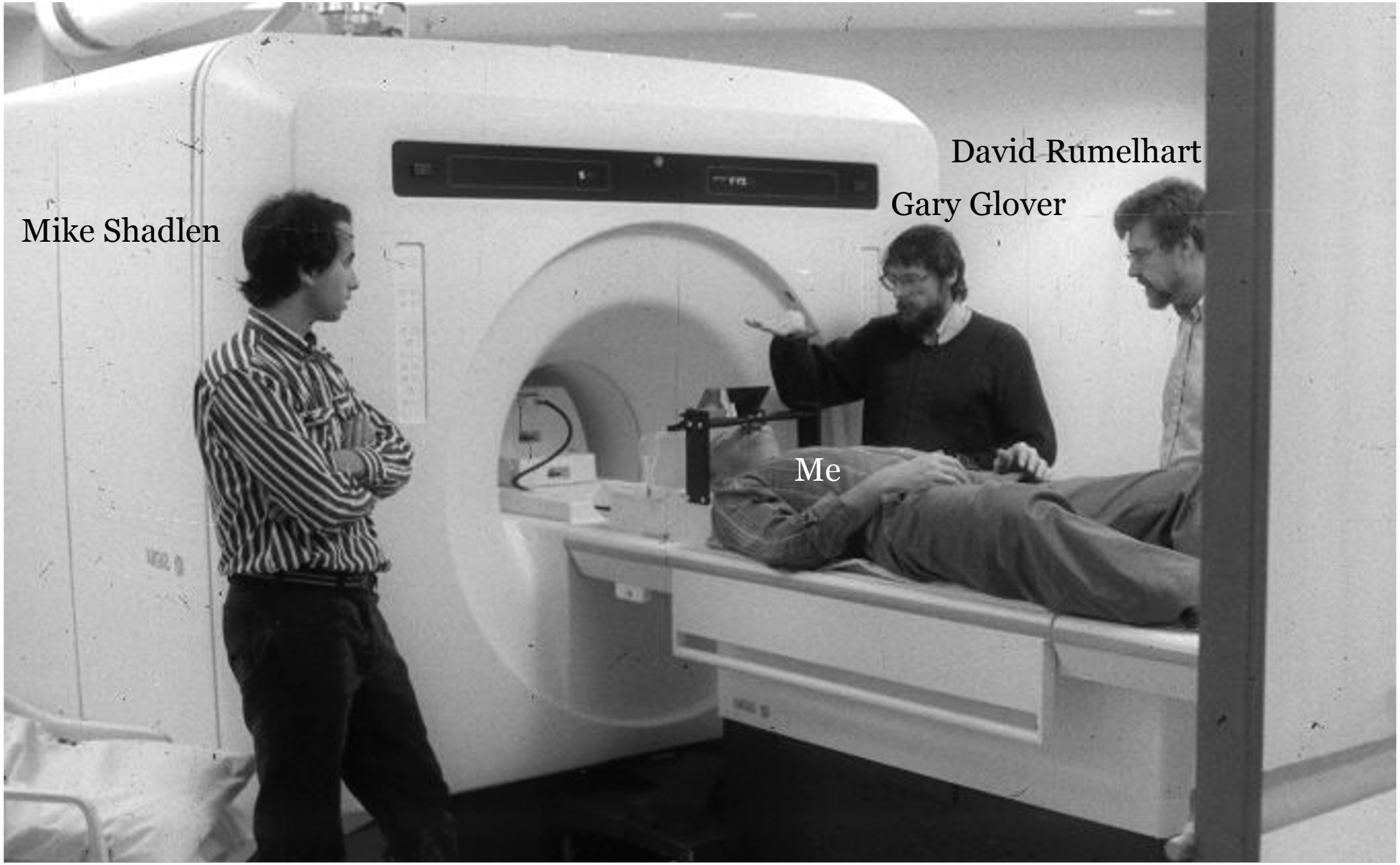
- The cortex is a sheet (2-4 mm thick) of tissue that covers the surface of the brain; other subcortical regions and types of cells matter too!

- Neurons/mm^3: 10^4-10^6
- Cortical Neurons: 10^{11}
- Synapses/neuron: 10^3
- Cortical Synapses: 10^{14}
- Surface area of each hemisphere: 20 x 30 cm^2

image from Graham Johnson

Neuron: impulse-conducting cell; bodies are in the cerebral cortex
Axon: a thin fiber that carries the output impulses from a neuron
Dendrite: a branching process of a neuron that receives impulses from other neurons
Synapse: The point of connection between neurons

VISTA Vision Science: MRI – September, 1991

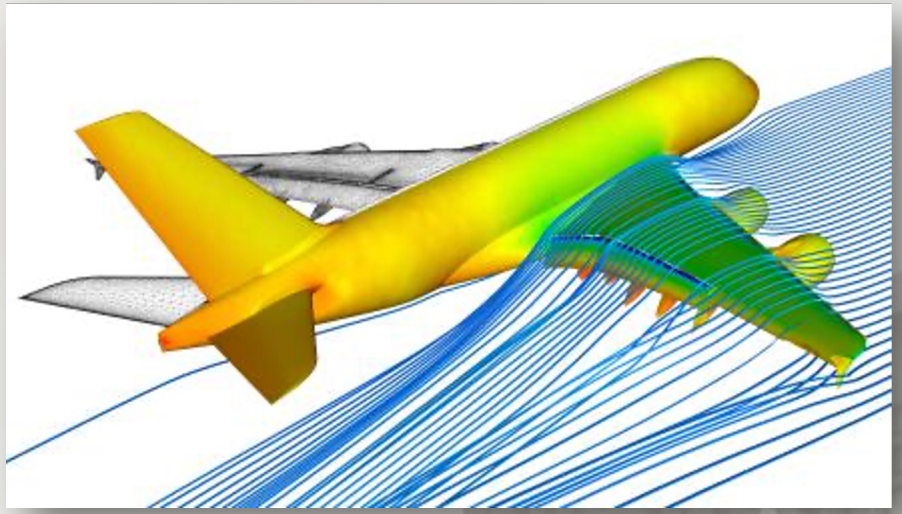


Founding director of the MRI center across the street in Building 420

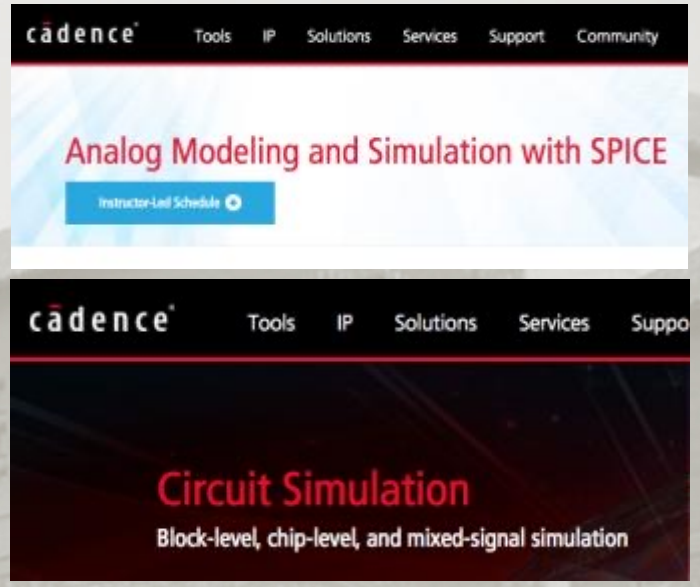
Simulation systems simulation is important in many mature industries

Image systems have become mature

ECU (Electronic Control Unit) Simulation for Automobiles



Numerical flow simulation on an Airbus A380



Integrated circuitry

Simulation is particularly useful for industry innovations

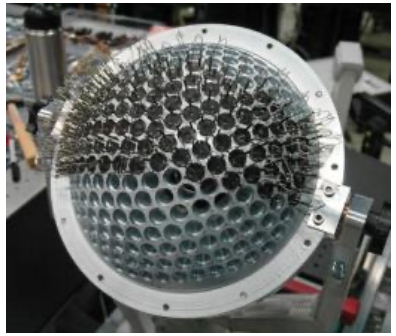
- Informal simulations occur routinely in the imaging industry
- Software development is usually custom, in-house

Multiple lens

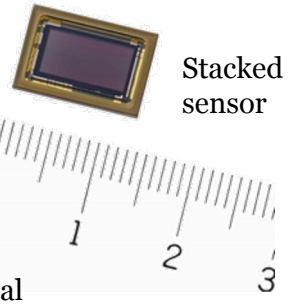
RGB-depth

Light field

360 Surround Video

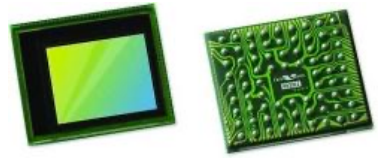


Massive resolution

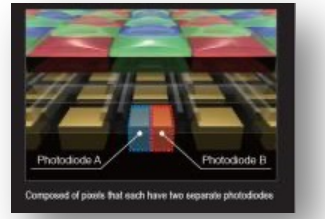


Stacked sensor

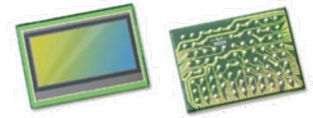
Global shutter



Dual pixel autofocus



RCCC automotive



Current status: Simulation of imaging components

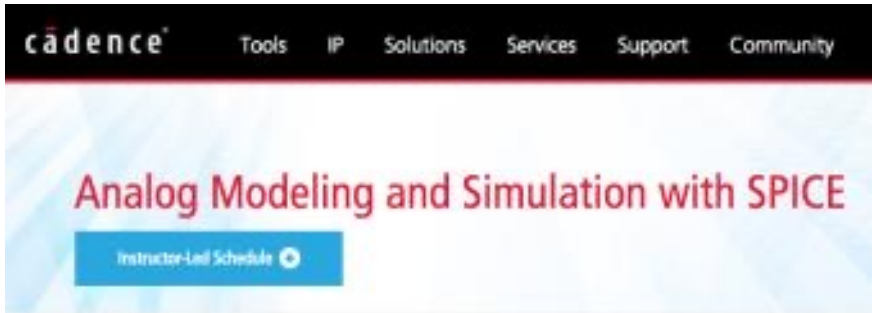
Lens design

Silicon design and Photonic Structures

SYNOPSYS®
Silicon to Software™

lumerical

Integrated circuitry



2003: We developed a 2D imaging systems simulation (physical units)

Image Systems Engineering Toolbox for cameras (**ISETCam**)

- End-to-end simulation (radiance to sensor)
  - Physical units (photons to electrons)

Optics

Sensor

Display

Imaging Systems Engineering Toolbox (ISETCam)

More than 500 users in
80 companies,
9 research institutes,
65 universities,
in 24 countries

Open Sourced on GitHub
in 2018

2018: We extended to 3D simulations and larger scale compute for ML

More about this later

Asset generation

blender

Scene Auto-assembly

FLYWHEEL

Scene spectral irradiance rendering

ISET3D

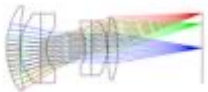
PBRT (docker)

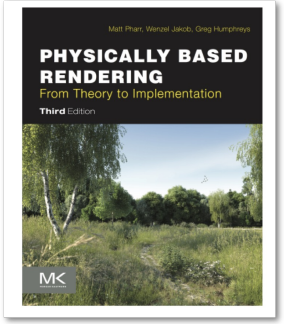
(q/s/sr/nm/m²)

Sensor irradiance

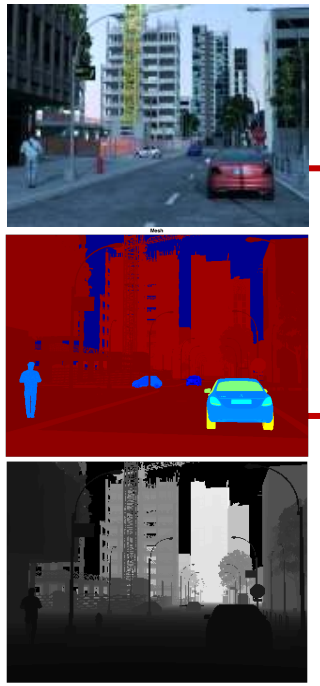
(q/s/nm/m²)

Pixel level depth and labels

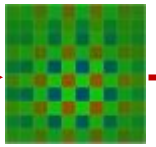




Sensor modeling



(electrons/pixel)



ISP

Pixel-level labeled sensor data

Annotation

ISETCAM

Training

PyTorch

TensorFlow

First point

Image systems simulation software that is trusted by key stakeholders in industry and academia can speed the development of next generation image sensors, camera arrays and displays.

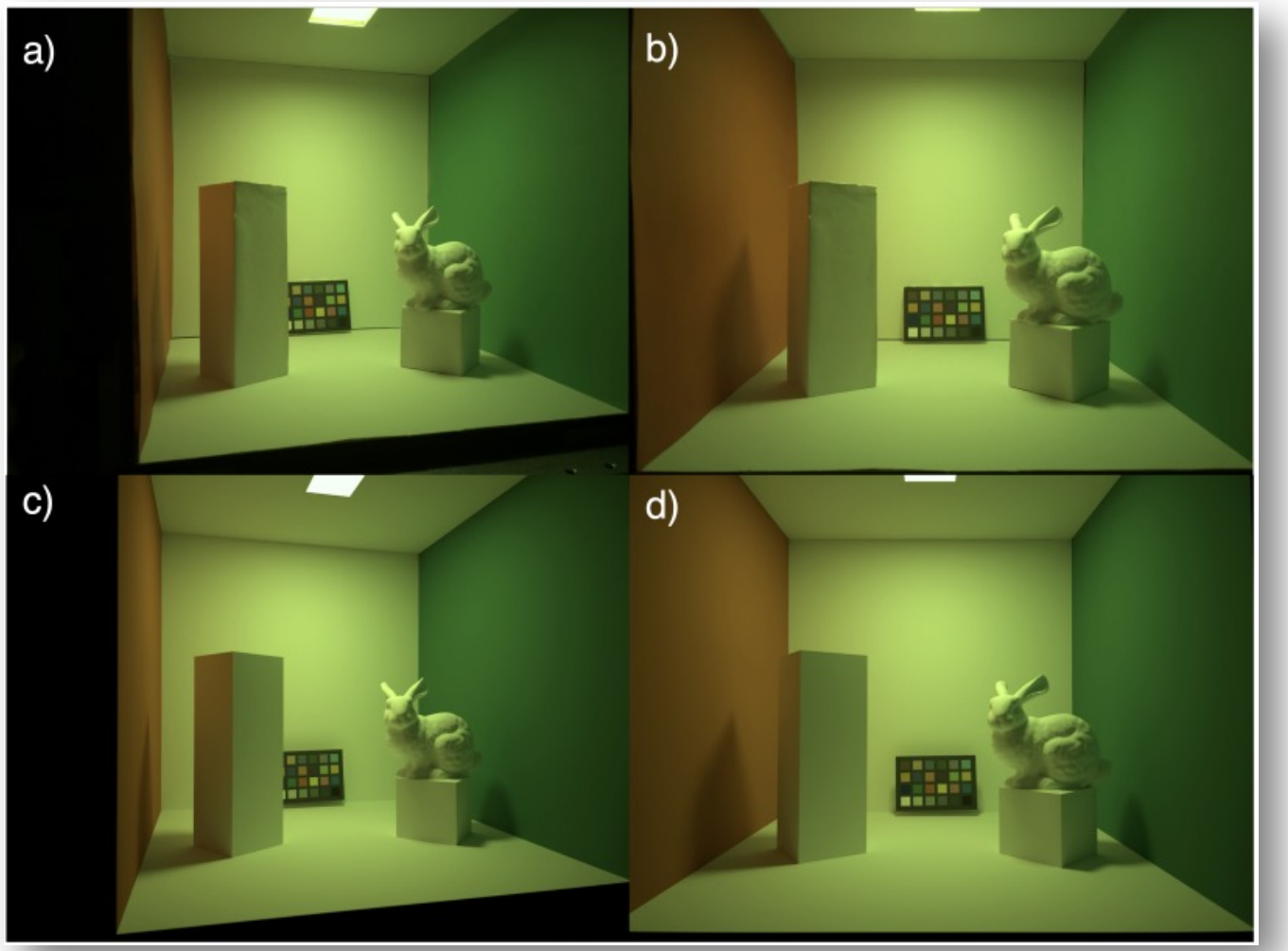
Image systems engineering tools (ISET)

- **ISET3d** is a Matlab toolbox that uses **PBRT** to calculate three-dimensional scene spectral radiance and sensor irradiance of complex scenes

- **ISETCam** is a Matlab toolbox that computes the image sensor response from the sensor irradiance; it also includes many industrial tools for evaluation of color (CIE standards) and spatial resolution (ISO standards)

Presentation outline

Part 1: Validation



Part 2: Automotive example

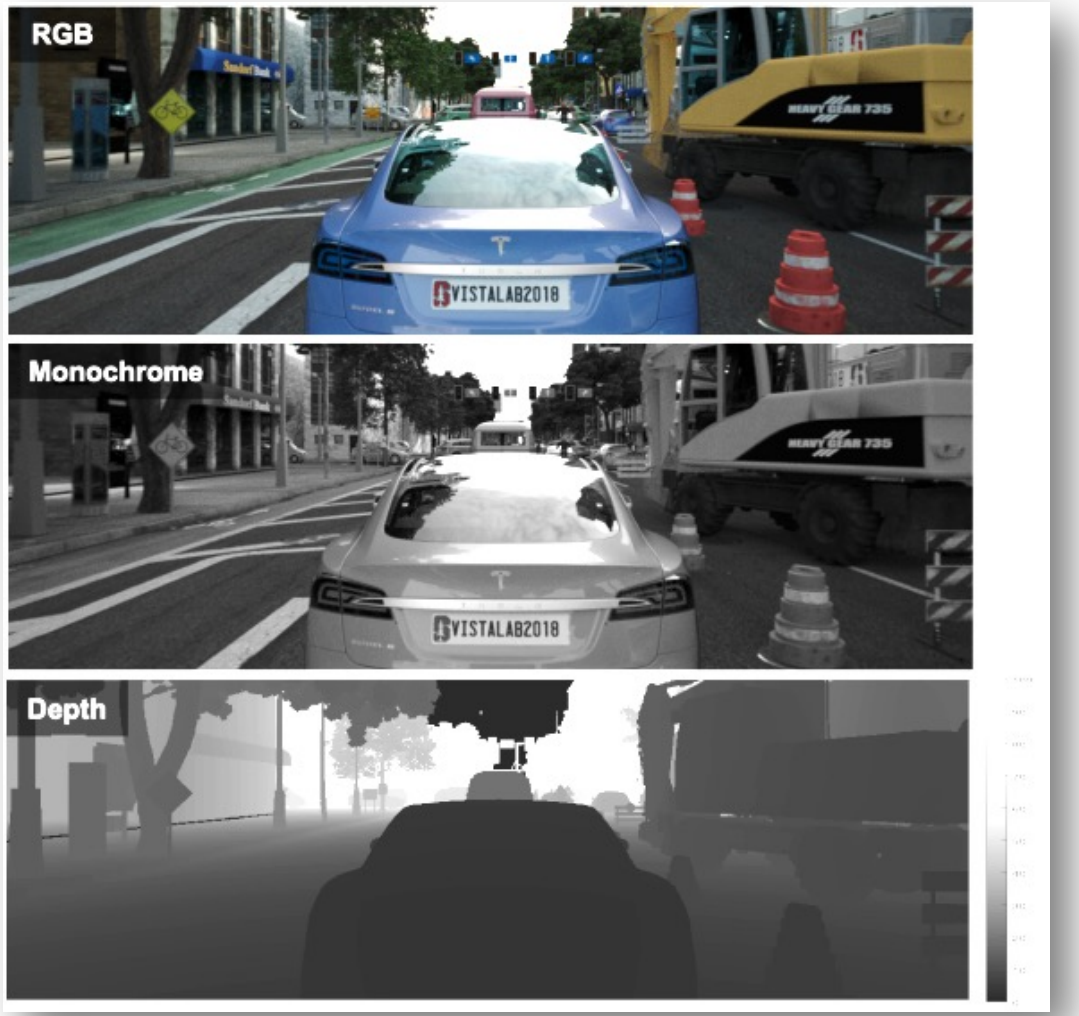
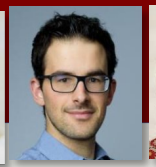


Image systems engineering tools validation

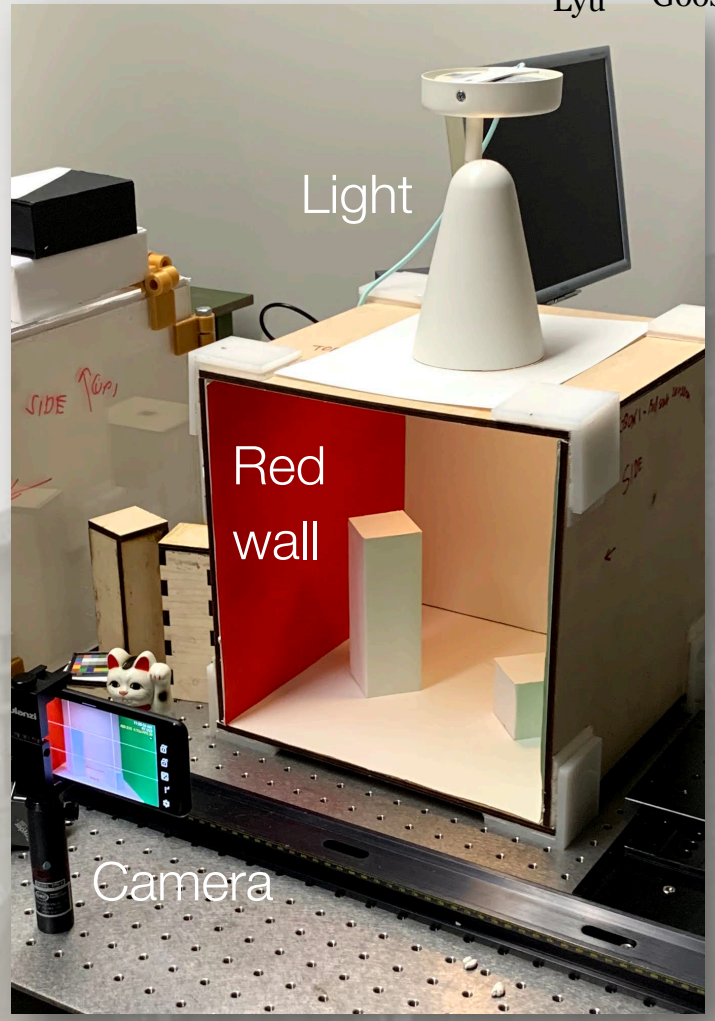

Zheng Lyu


Thomas Goossens


Joyce Farrell

- If you are designing or evaluating parts (optics, sensors), you must have accurate information

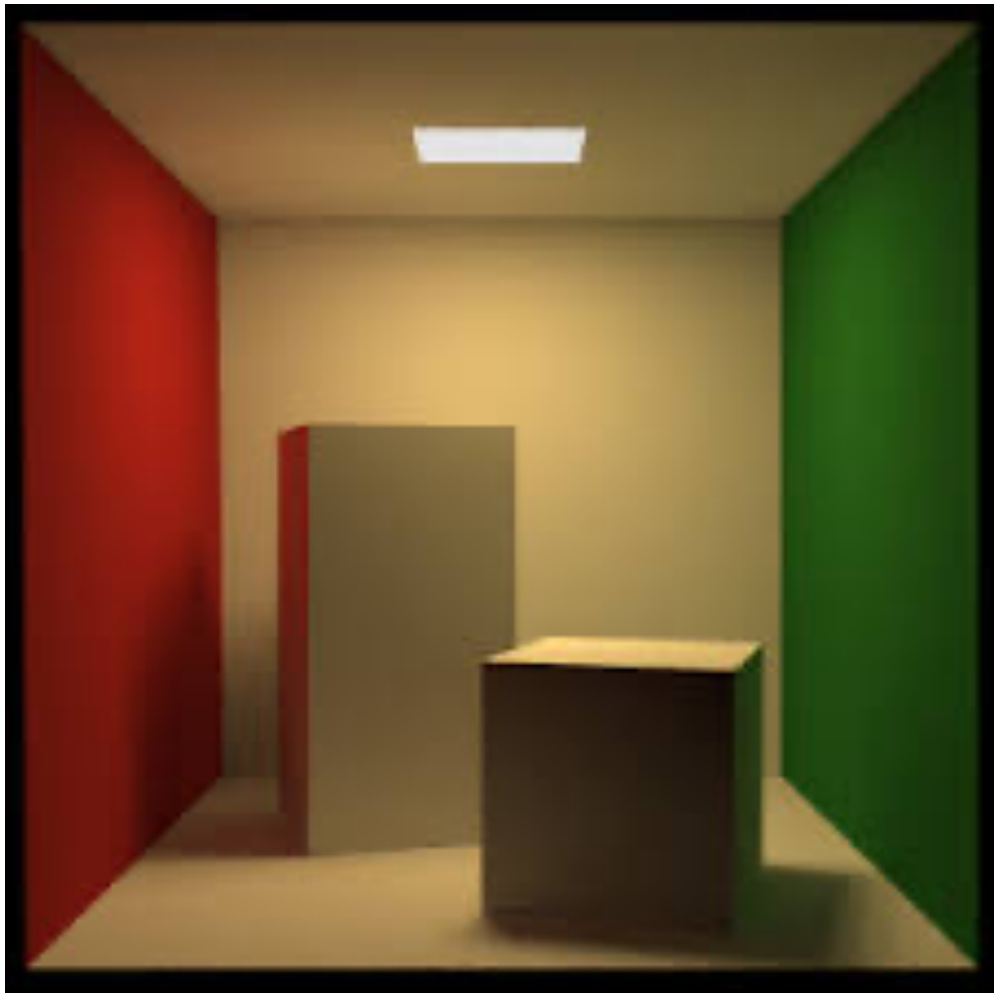
- If you are training a neural network, the camera properties matter for generalization



Accurate smartphone camera simulation using 3D scenes (2021.
Zheng Lyu, Thomas Goossens, Brian Wandell and Joyce Farrell
https://arxiv.org/abs/2201.07411

Validation tests: Cornell Box Construction

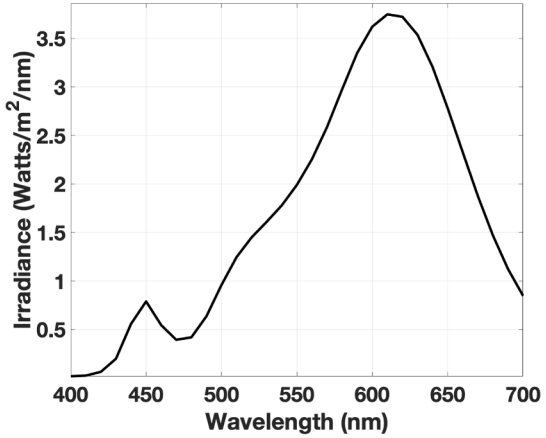
- The Cornell Box was developed to qualitatively test the **accuracy of computer graphics rendering**
- We use it to quantitatively test end-to-end simulation of **3D scenes, optics, and image sensors**



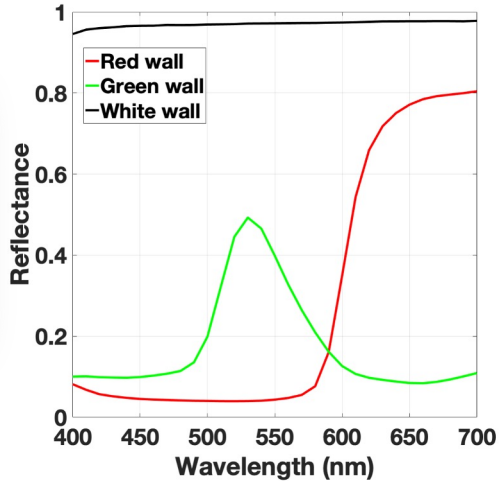
- Goral, Torrance, Greenberg, and Battaile, "Modeling the Interaction of Light Between Diffuse Surfaces," Computer Graphics (Proc. SIGGRAPH 84), Vol. 18, No. 3, July **1984**, pp. 213-222
- Meyer, Rushmeier, Cohen, Greenberg, and Torrance, "An Experimental Evaluation of Computer Graphics Imagery" ACM Transactions on Graphics, Vol. 5, No. 1, January **1986**, Pages 30-50.

We built one (the 3D scene)

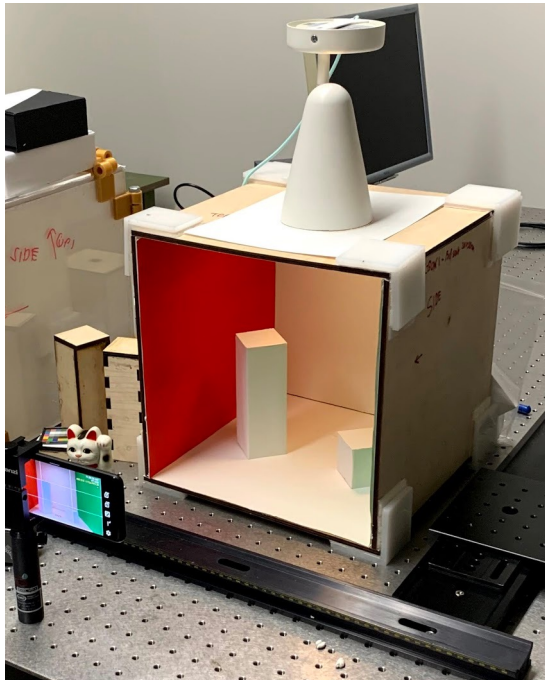
- We measured the asset sizes and scene geometry, surface reflectance, illuminant spectral power distribution

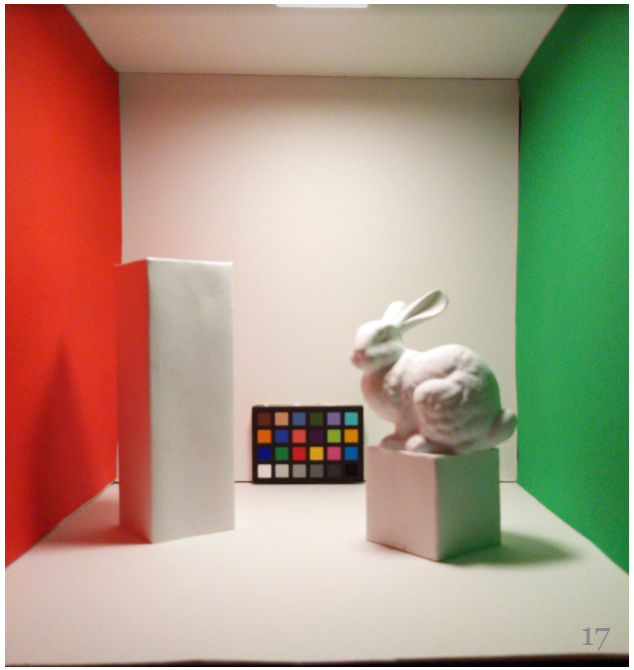


Illuminant SPD



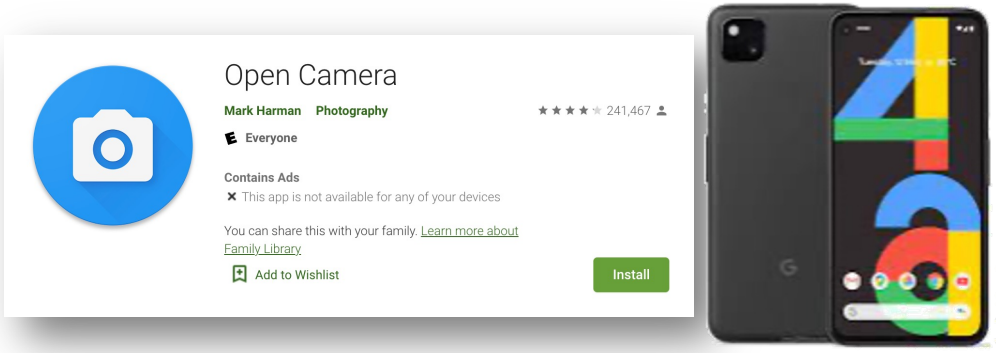
Surface spectral reflectance

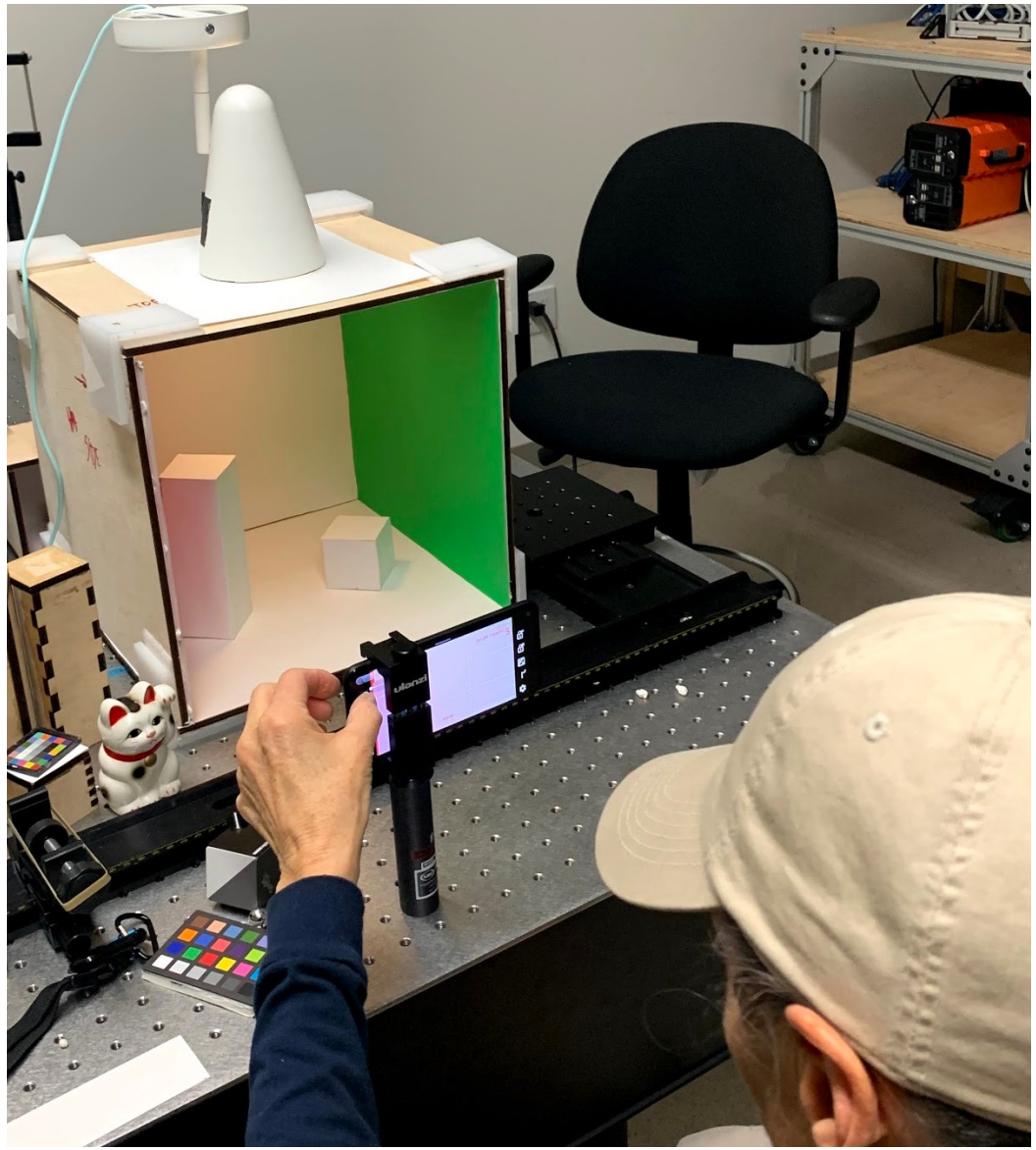




Sensor data from the camera

- We used a Google Pixel 4a and the OpenCamera app to acquire **relatively raw data**

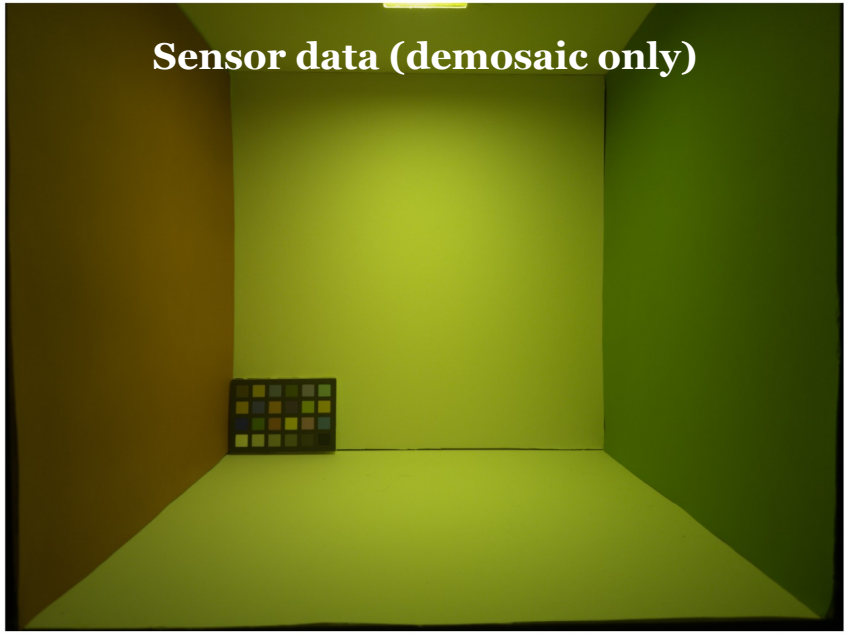


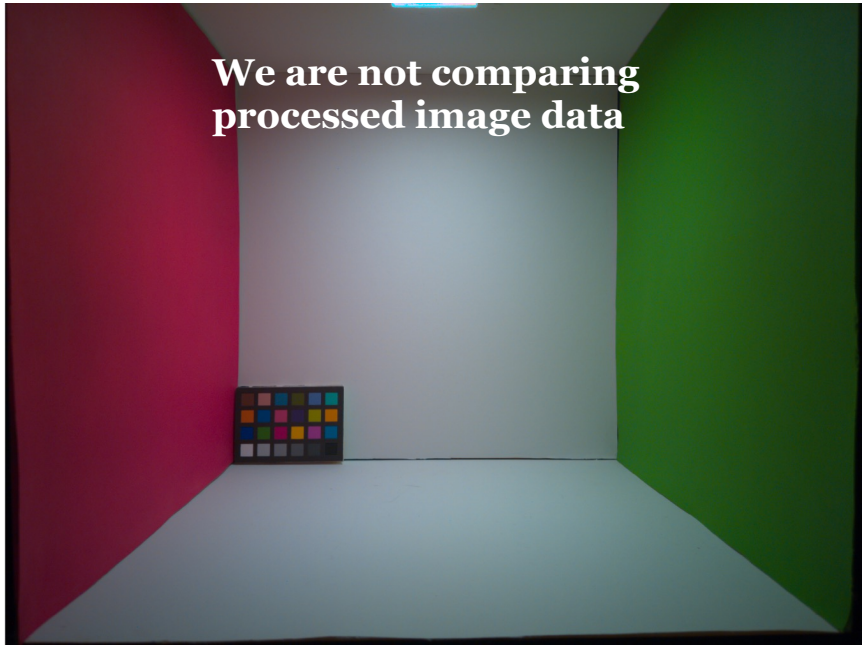


Sensor renderings

- Our goal is to match the **digital sensor values**, not the processed data

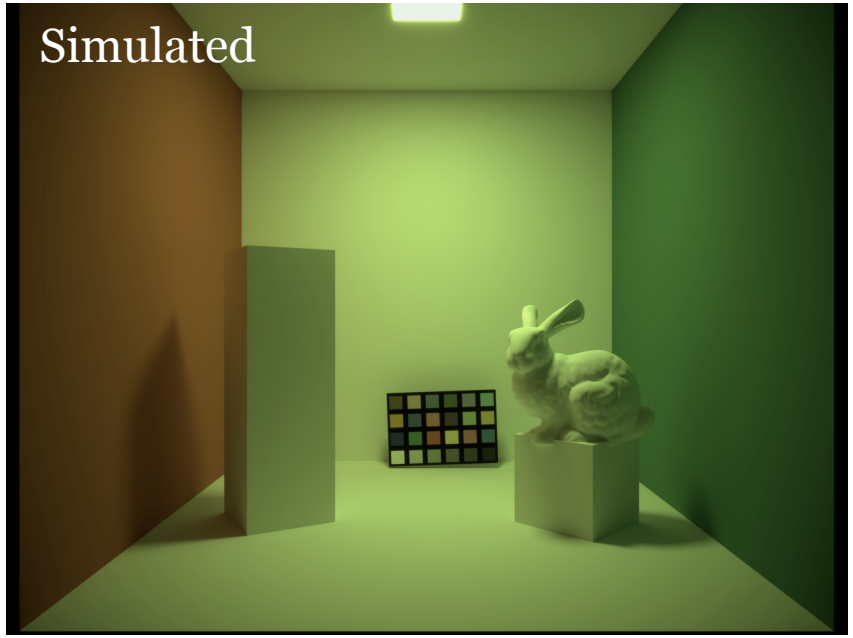
- The renderings I will show you are either **simulated or measured sensor data**, with only bilinear demosaicking



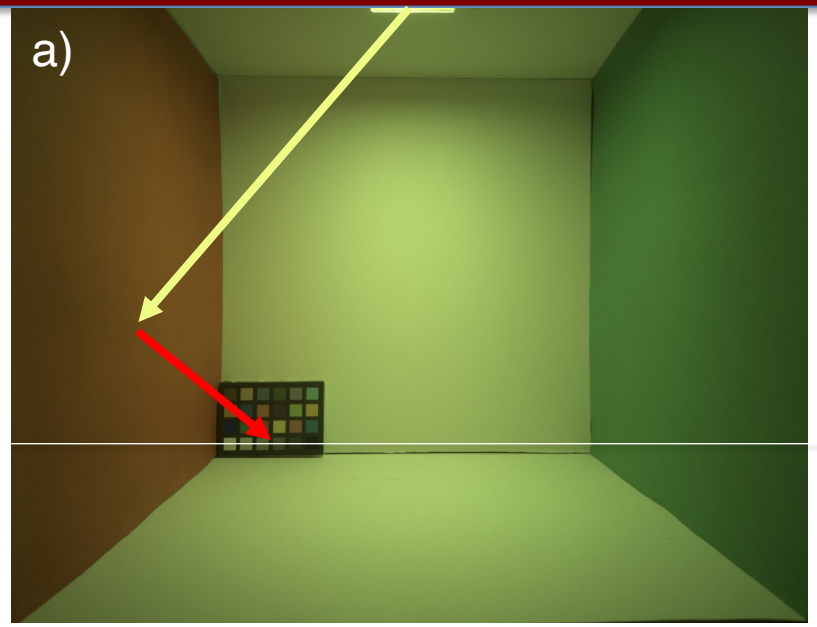


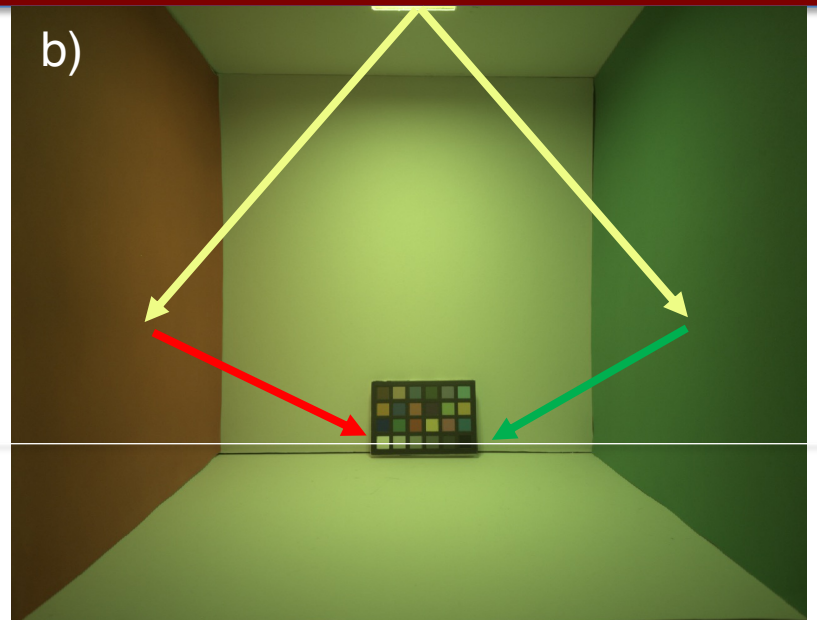
Quick check based just on appearance

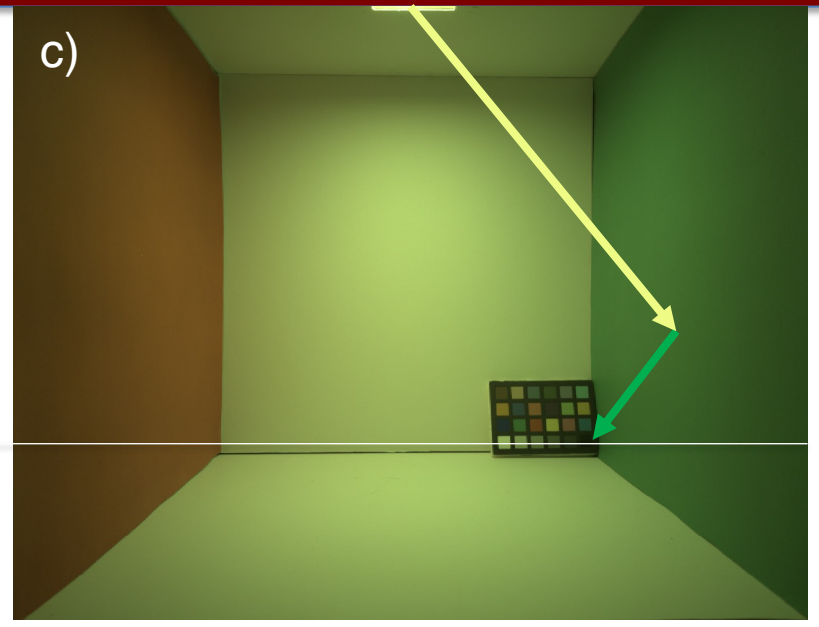
- Overall similarity is good
- High dynamic range
- Shadows
- Color interreflections (sides of the box reflect light from the wall)
- If you are designing hardware, this check is far from adequate: quantify!

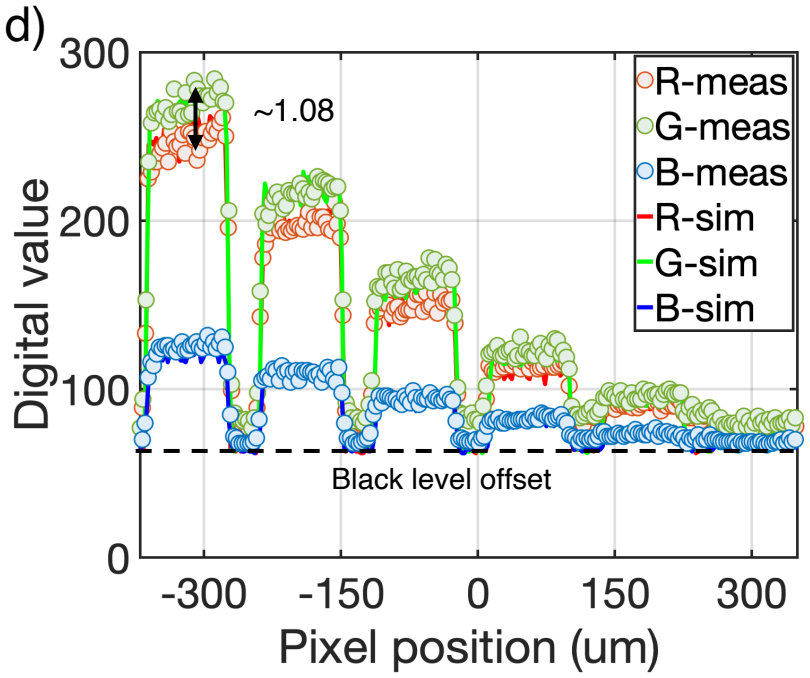


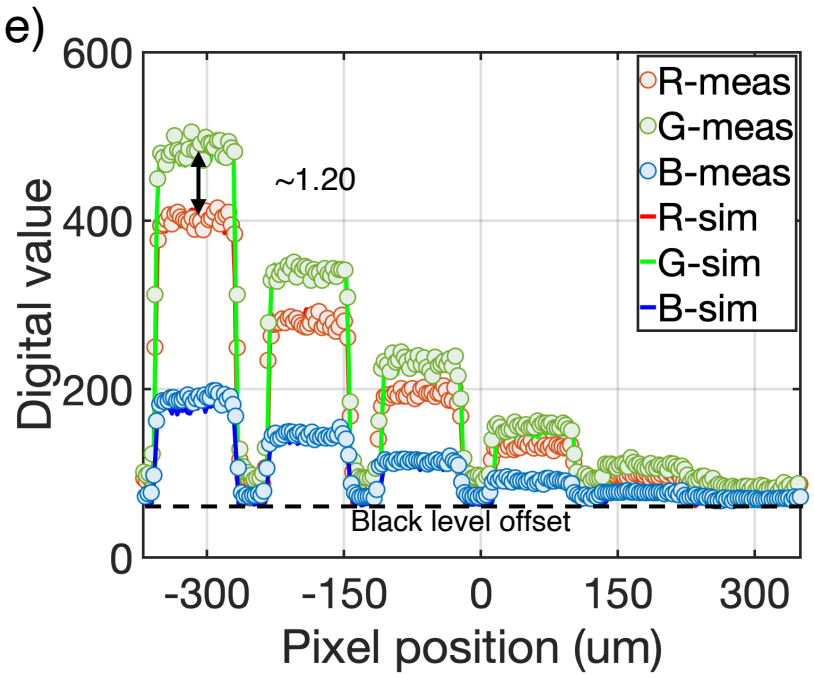
Surface inter-reflections

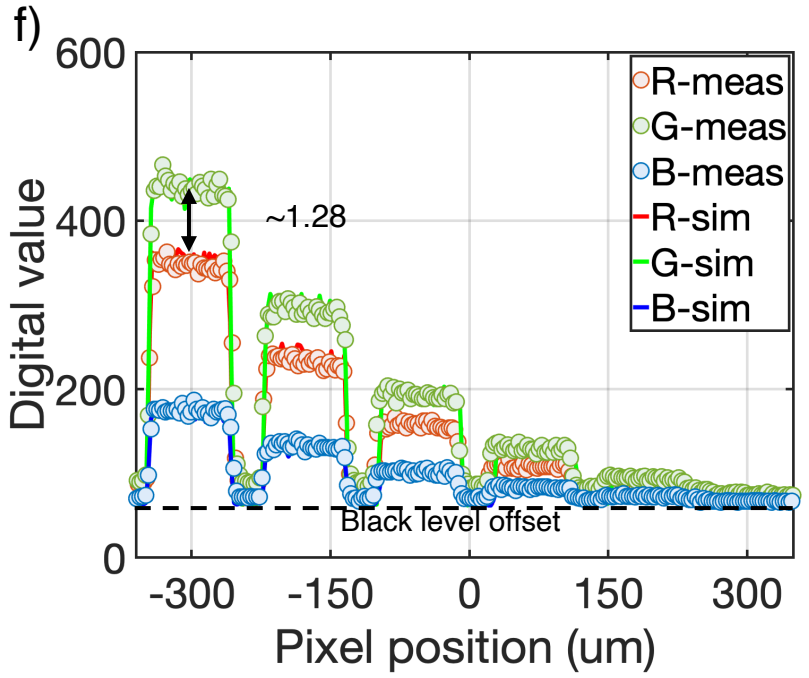






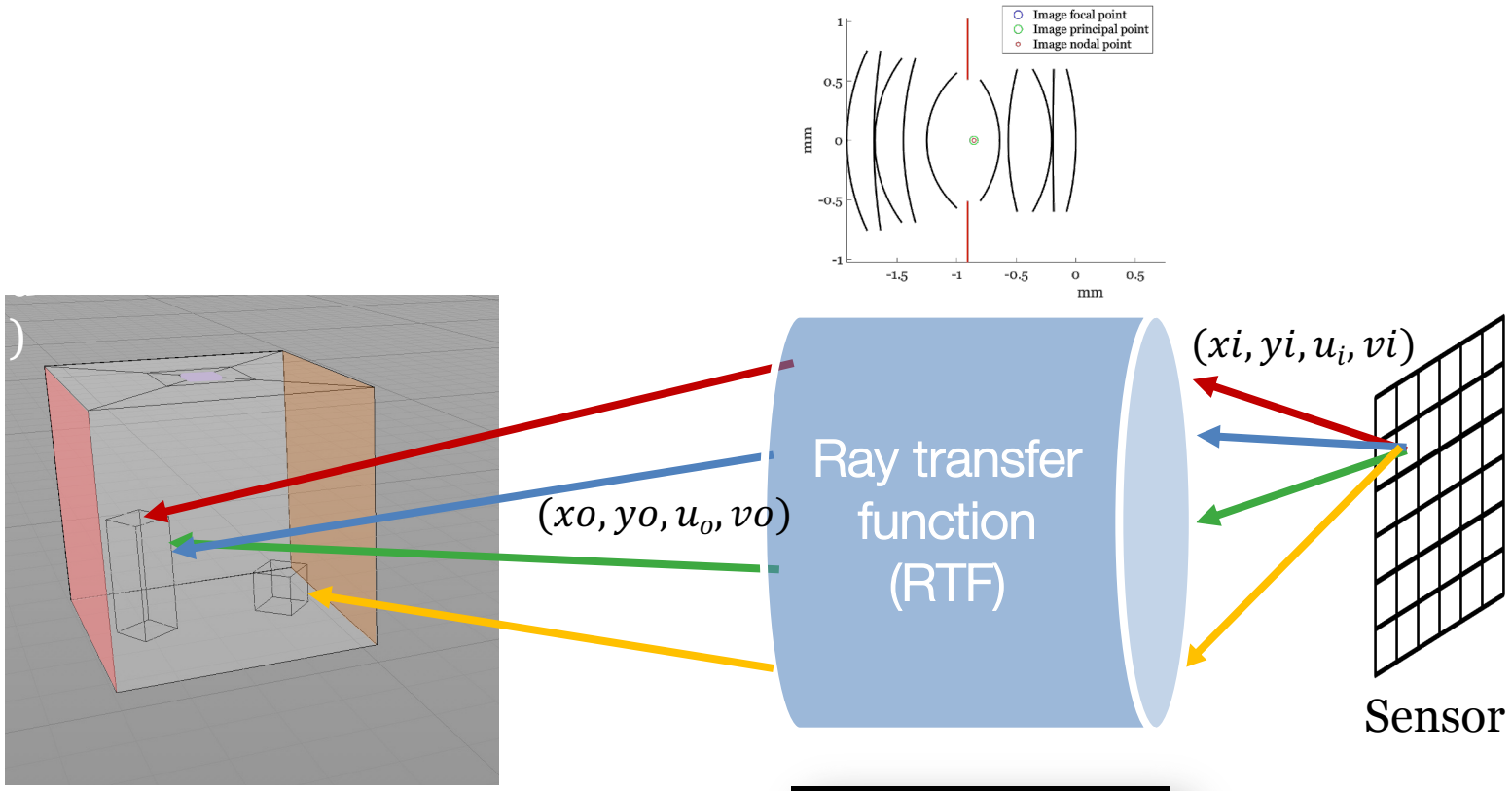






Optics: Ray trace function

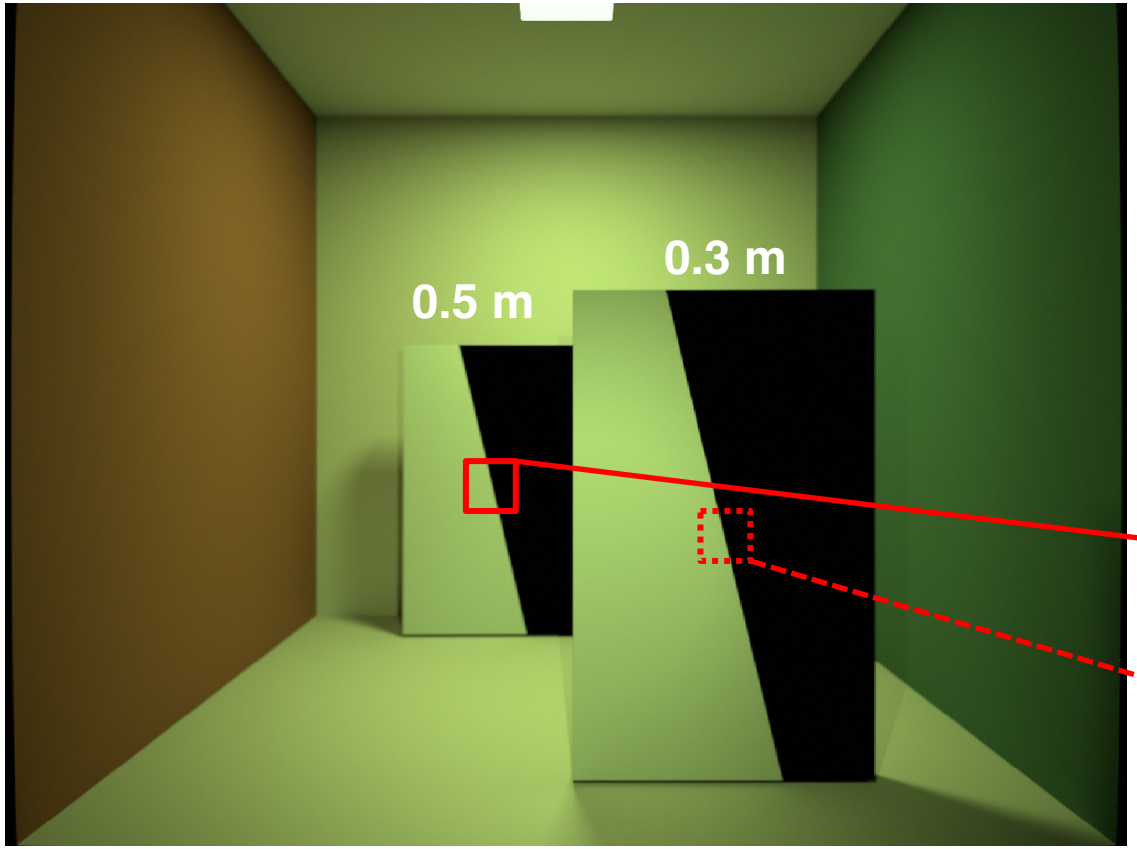
- Proprietary lens prescription
- Zemax black box model was made available
- TG developed and implemented a ray transfer function model in PBRT: a method to map input rays to output rays when the prescription is unknown

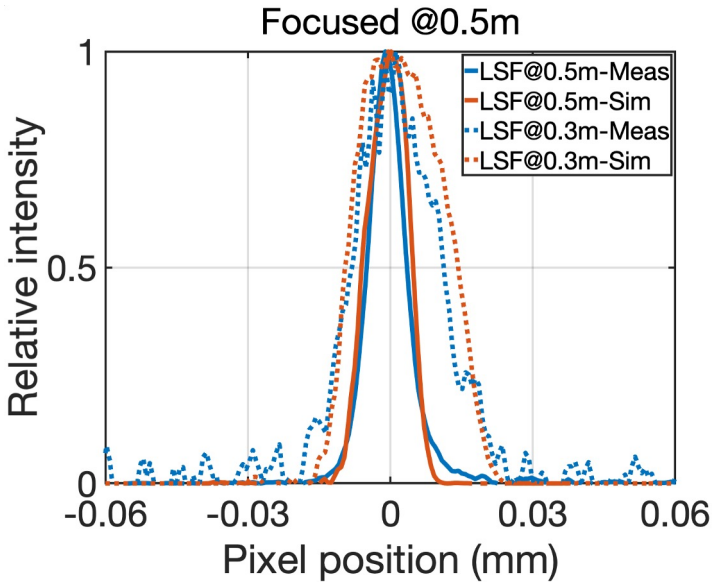


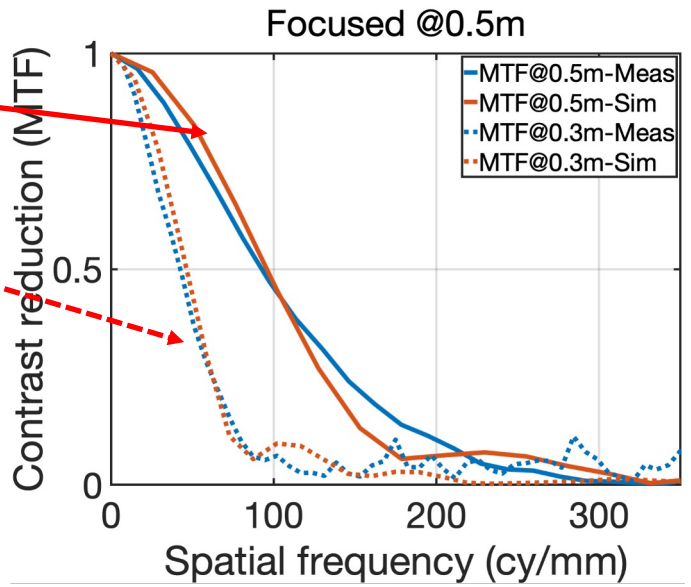
Ray-transfer function for modeling optics in camera simulation with 3D scenes (2022). Thomas Goossens, Zheng Lyu, Jamyuen Ko, Gordon Wan, Ricardo Motta, Joyce Farrell and Brian Wandell

Focus-dependent LSF/MTF (depth of field)

Focused at 0.5 meters

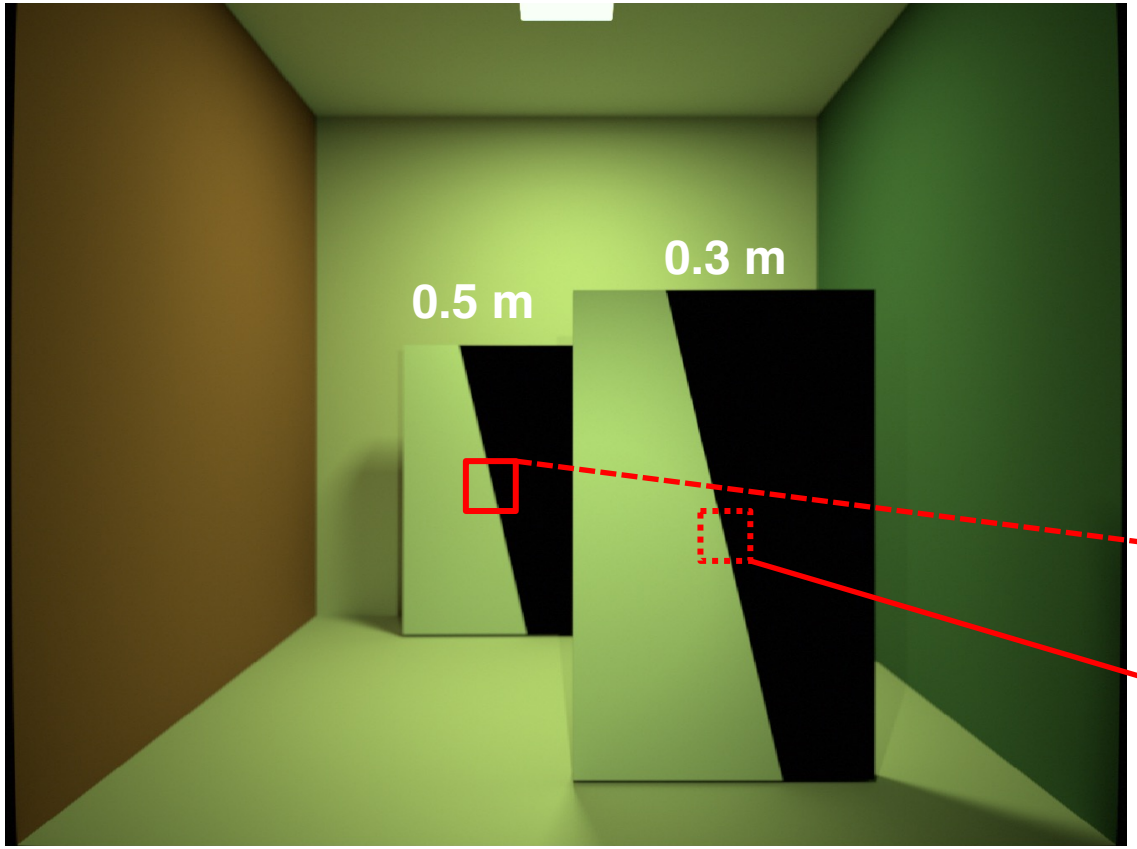


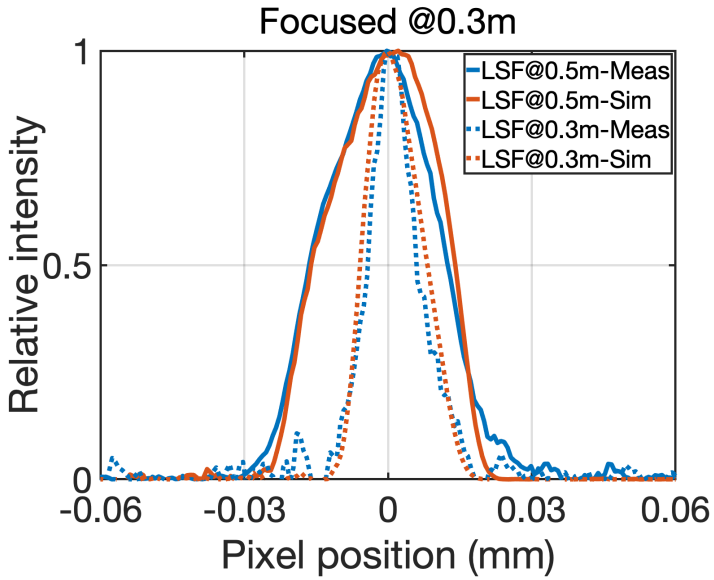


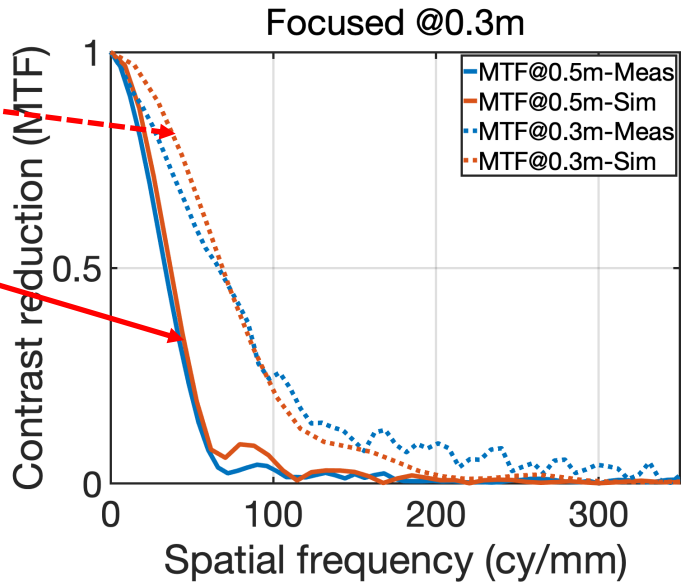


Focus-dependent LSF/MTF (depth of field)

Focused at 0.3 meters







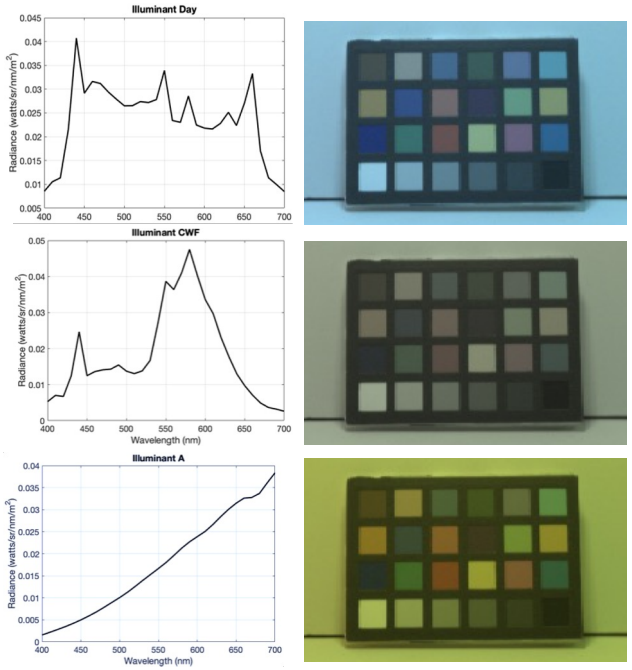
3D simulation system: sensor

- **Sensor**: The Google Pixel 4a uses the Sony IMX363 CMOS sensor
- **Electronics**: We used vendor data and also measured electronic noise properties, including read noise, dark noise, DSNU and PRNU

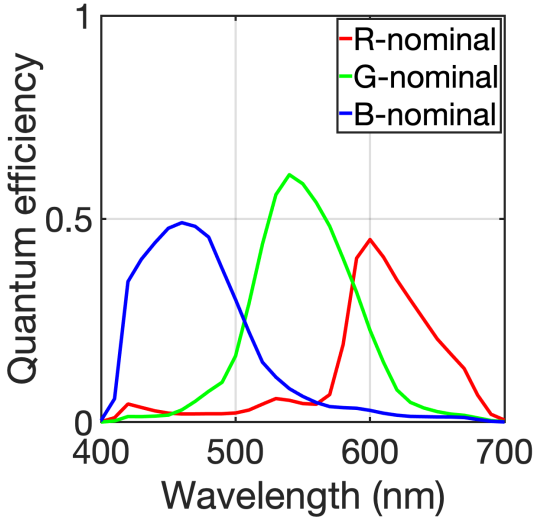
Properties	Parameters	Values (units)
Geometric	Pixel size	[1.4, 1.4] (um)
	Fill factor	100 (%)
Electronics	Well capacity	6000 (# e-)
	Voltage swing	0.4591 (volts)
	Conversion gain	7.65×10^{-5} (Volts/e-)
	Analog gain	1
	Analog offset	0.0287
	Quantization method	10 bits
Noise sources @ Analog gain = 1	DSNU	0.64 (mV)
	PRNU	1.9 （%）
	Dark voltage	0
	Read noise	5 (mV)

Color channel quantum efficiencies

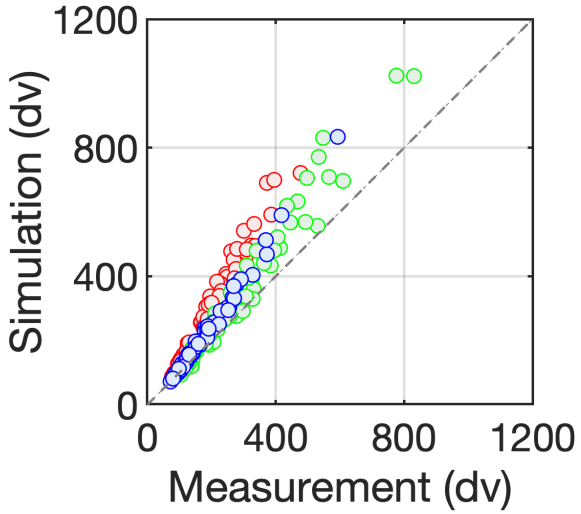
24 color patches
3 illuminants



Vendor specification

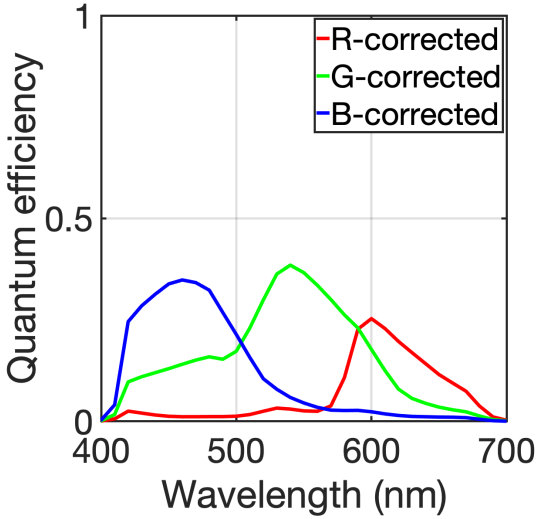


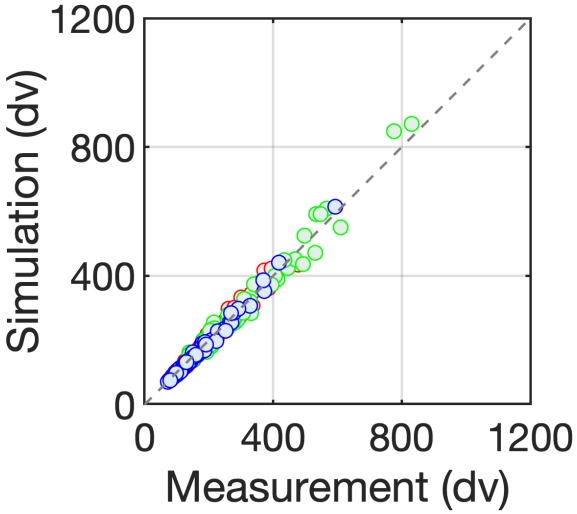
Calibrated correction: color channel gain and crosstalk



$$(r \quad g \quad b)\begin{pmatrix} .56 & .08 & .01 \\ 0 & .59 & 0 \\ 0 & .25 & .71 \end{pmatrix}$$

Calibration

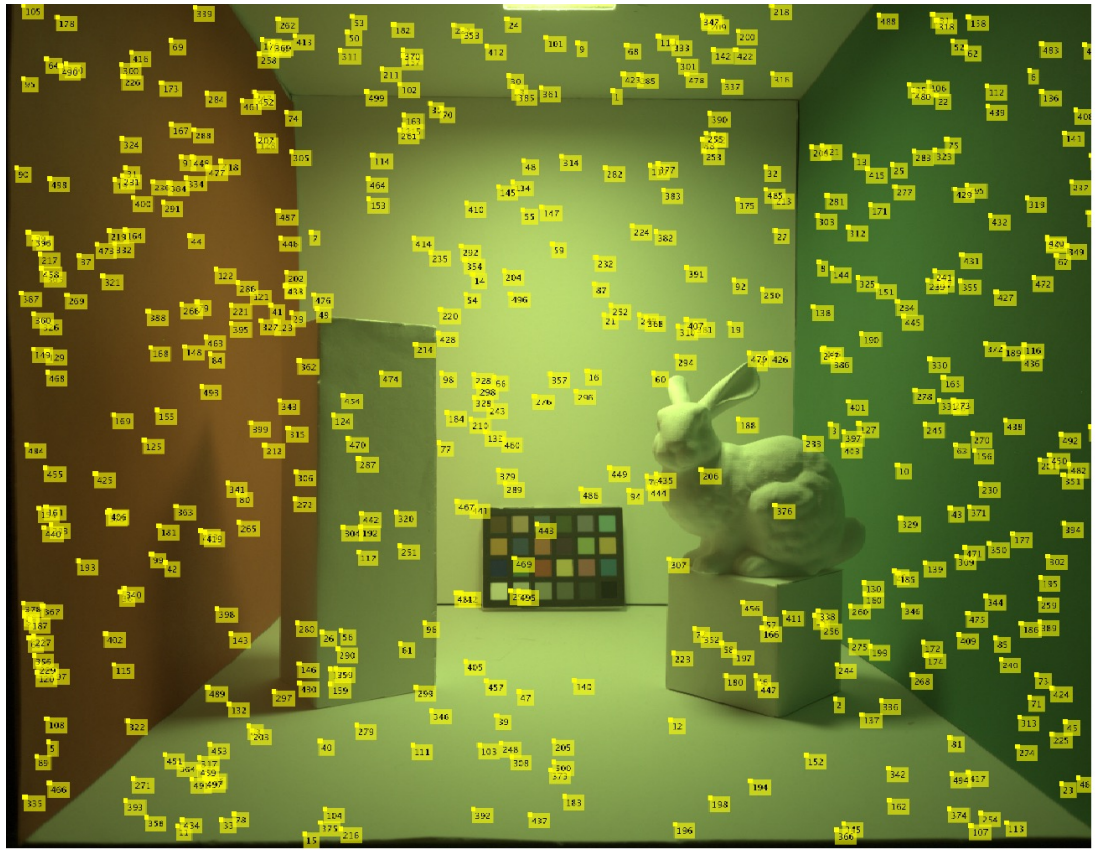




Sensor noise measurements

A combination of electronic noise and photon noise

Many small uniform regions



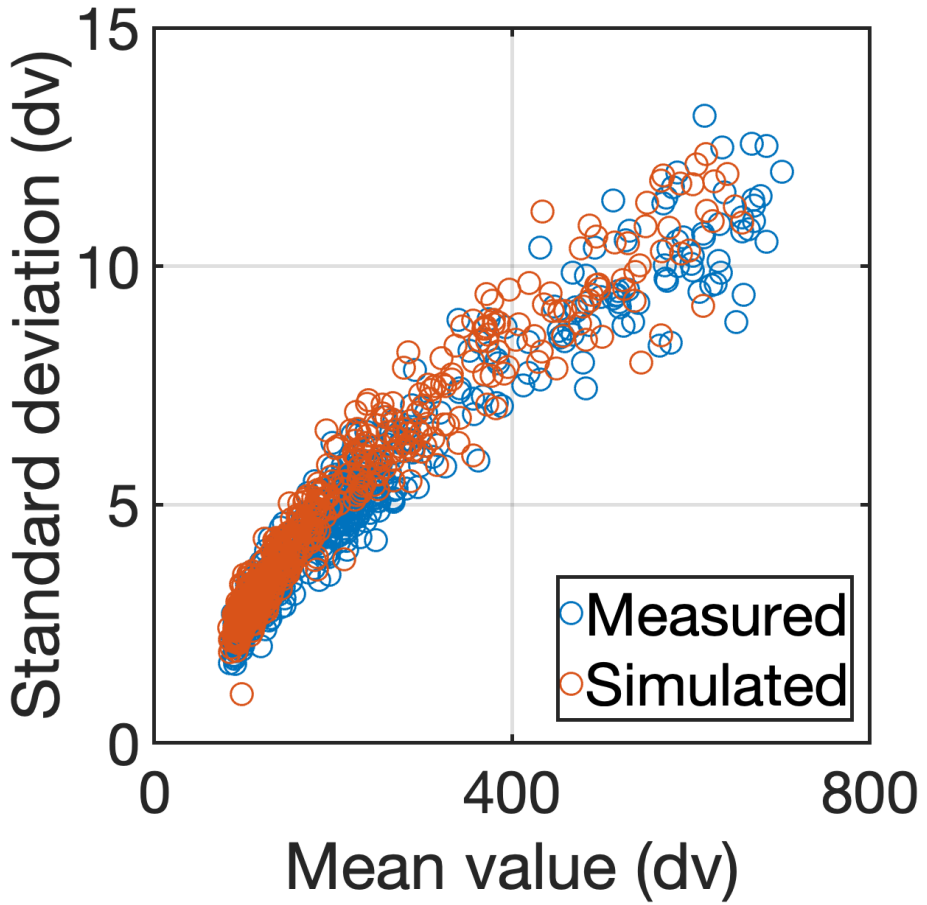
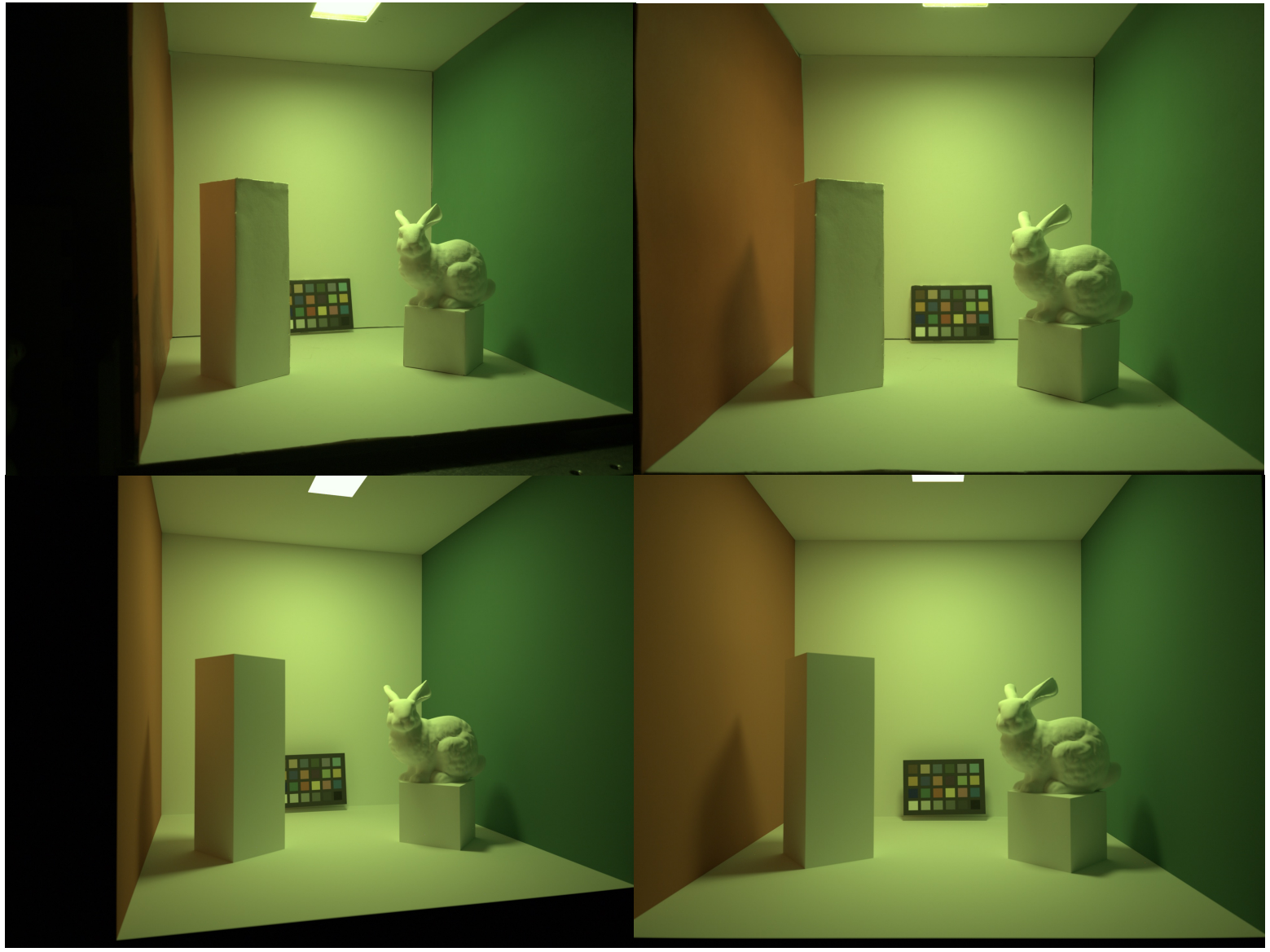


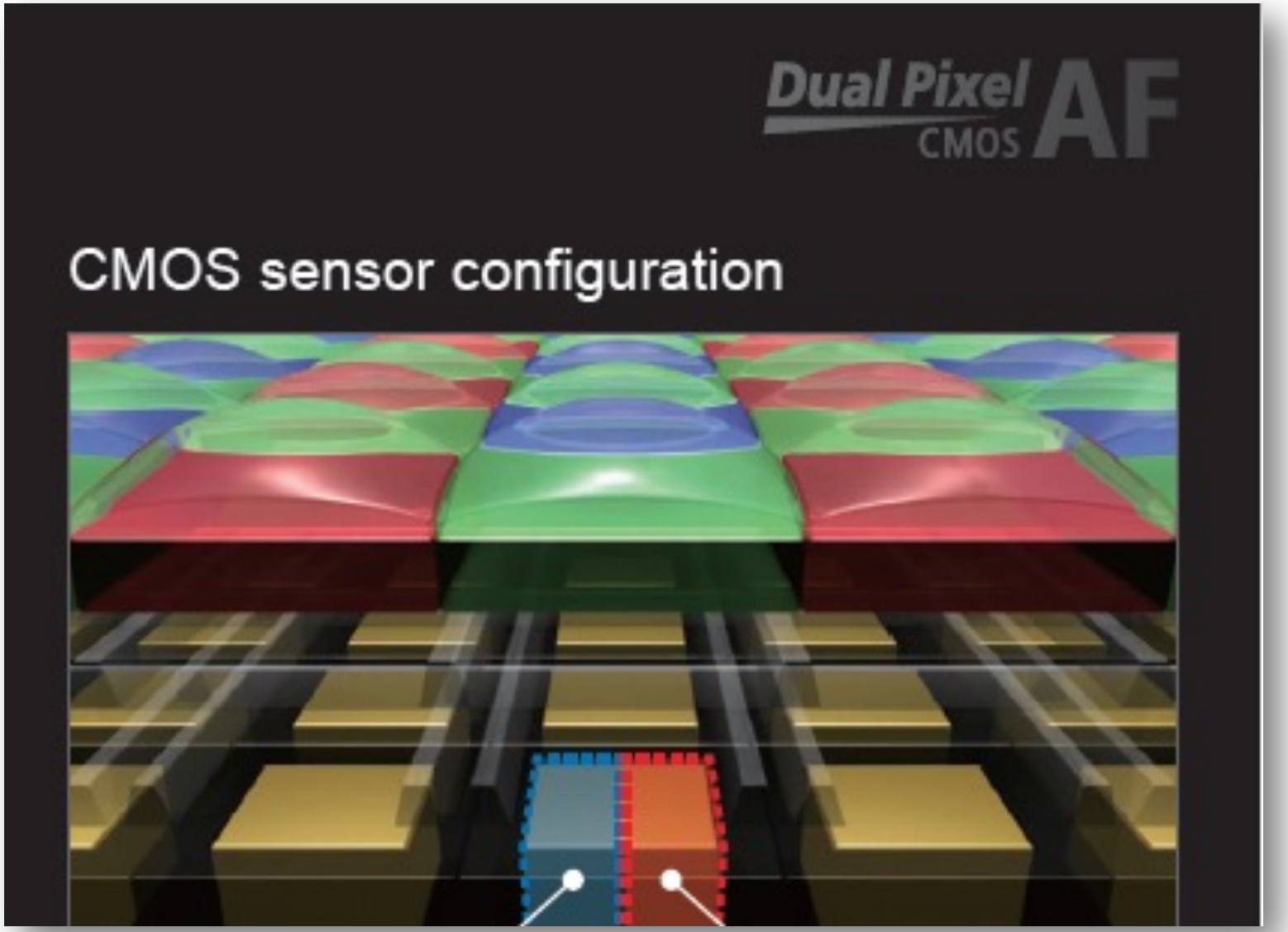
Image system simulation evaluation using 3D scenes

- ✓ Surface inter-reflections
- ✓ Depth of focus
- ✓ Vignetting
- ✓ Sensor quantum efficiency
- ✓ Sensor noise



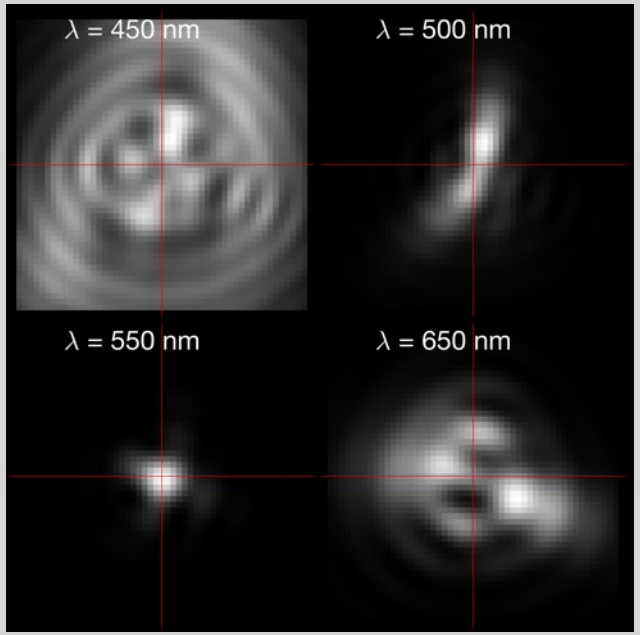
Next: Wave properties, light field designs, camera controllers

- Dual pixel (Canon) for autofocus
- Quad pixel and more for single-shot depth



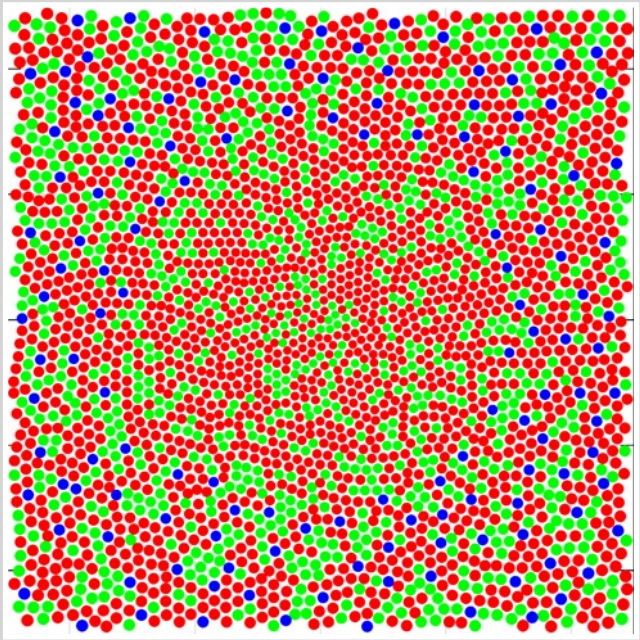
ISETBio: Accounting for absolute sensitivity: modern estimates of optics/mosaic

Wavefront aberration - based optics

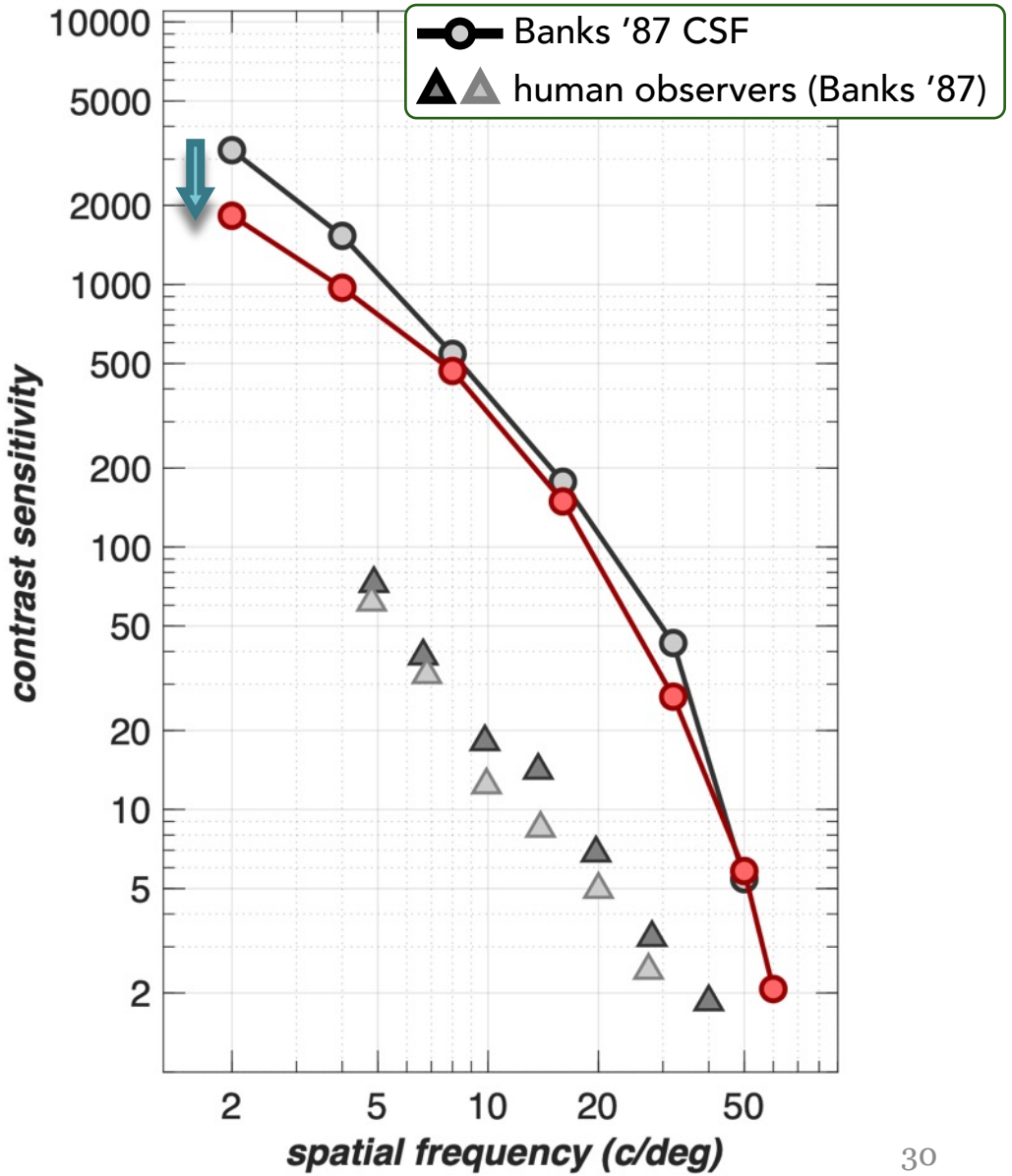


(from Thibos et al. '92)

Eccentricity-dependent density mosaics

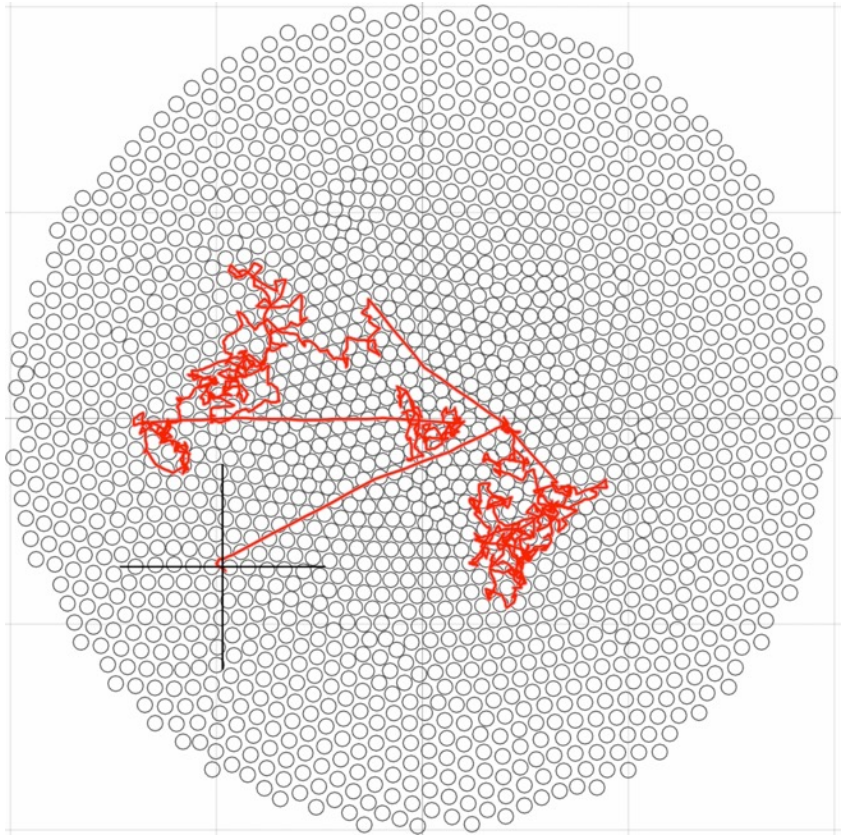


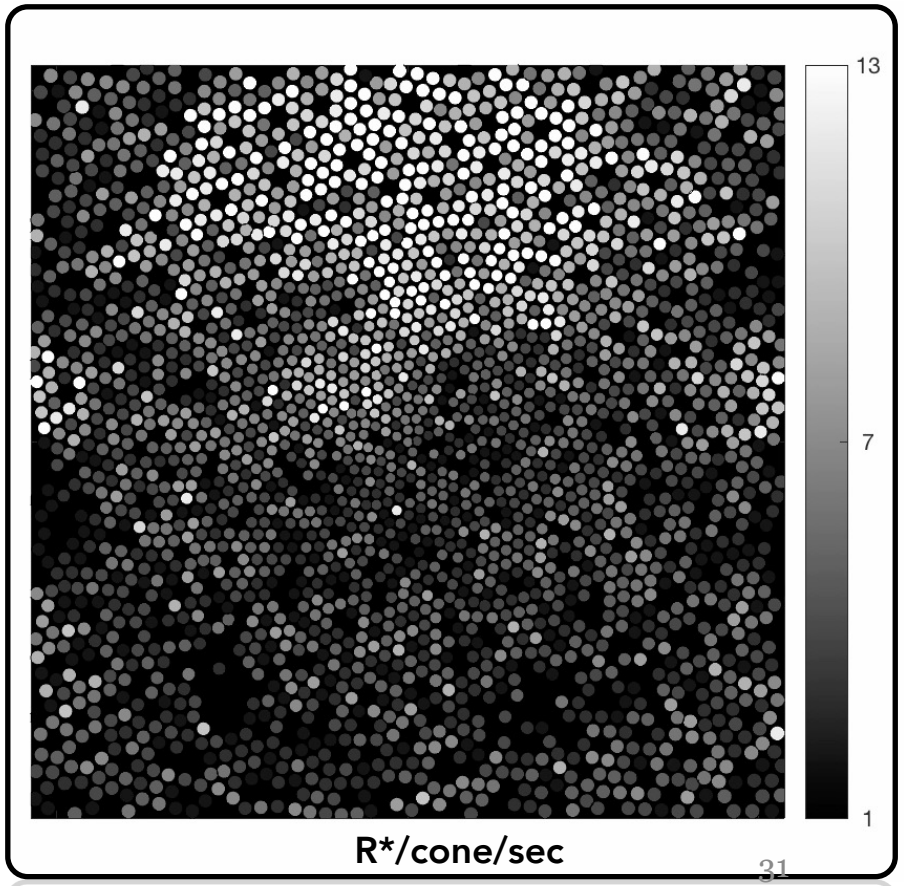
(based on Curcio et al. '90)



Accounting for absolute sensitivity: fixational drift (Engbert and Kliegl, 2004)

cone photopigment excitation → fixational eye movement object → cone photopigment excitation sequence

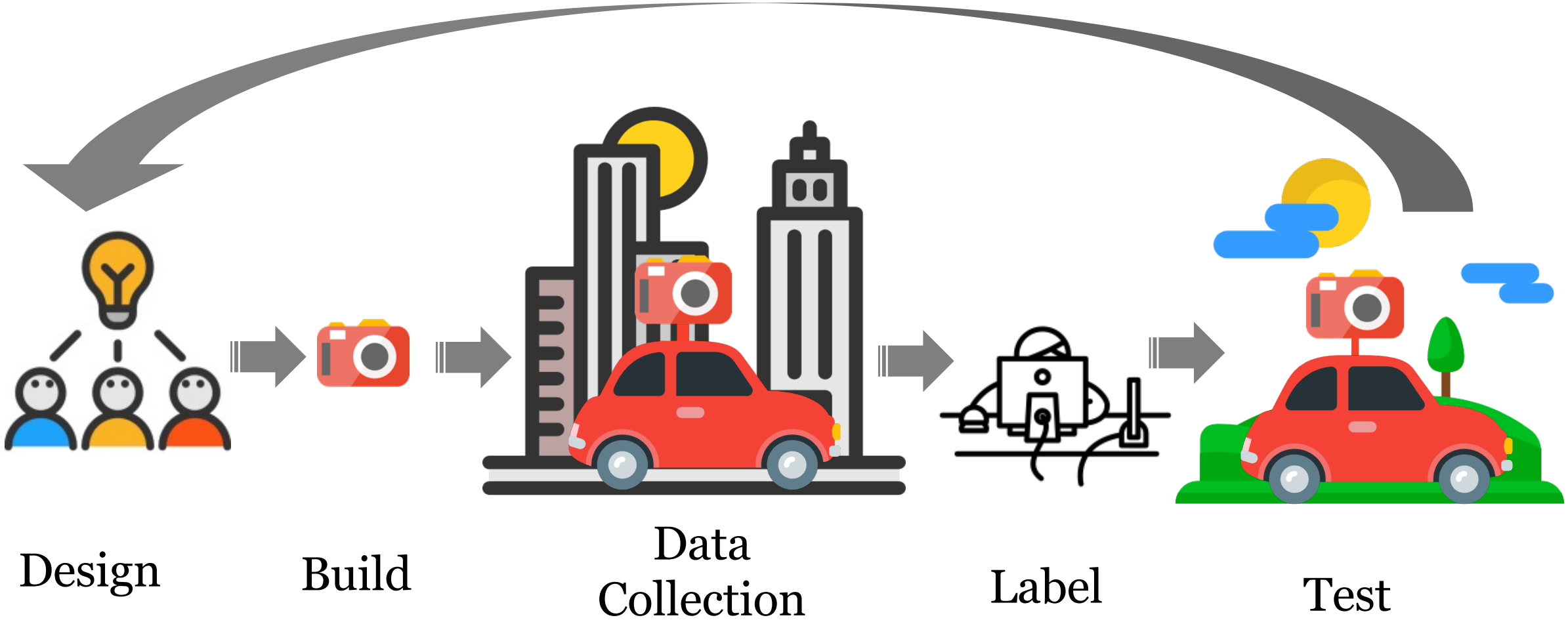




Camera design for a machine-learning application

- ISET3d, ISETCam
- Automotive cameras and ML

Camera design for automotive machine-learning



Synthetic camera data for automotive applications: We care about the camera!

Simulation system overview for autonomous driving

Asset generation

Glass Metal Carpaint

blender

Scene Auto-assembly

FLYWHEEL

Scene spectral irradiance rendering

ISET3D

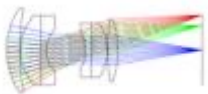
PBRT (docker)

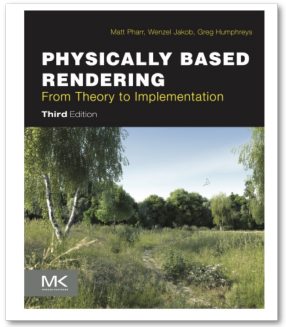
(q/s/sr/nm/m²)

Sensor irradiance

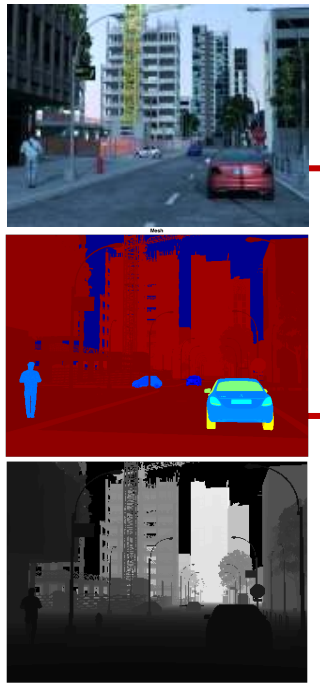
(q/s/nm/m²)

Pixel level depth and labels





Sensor modeling



(electrons/pixel)

ISP

Pixel-level labeled sensor data

Annotation

ISETCAM

Training

PyTorch

TensorFlow

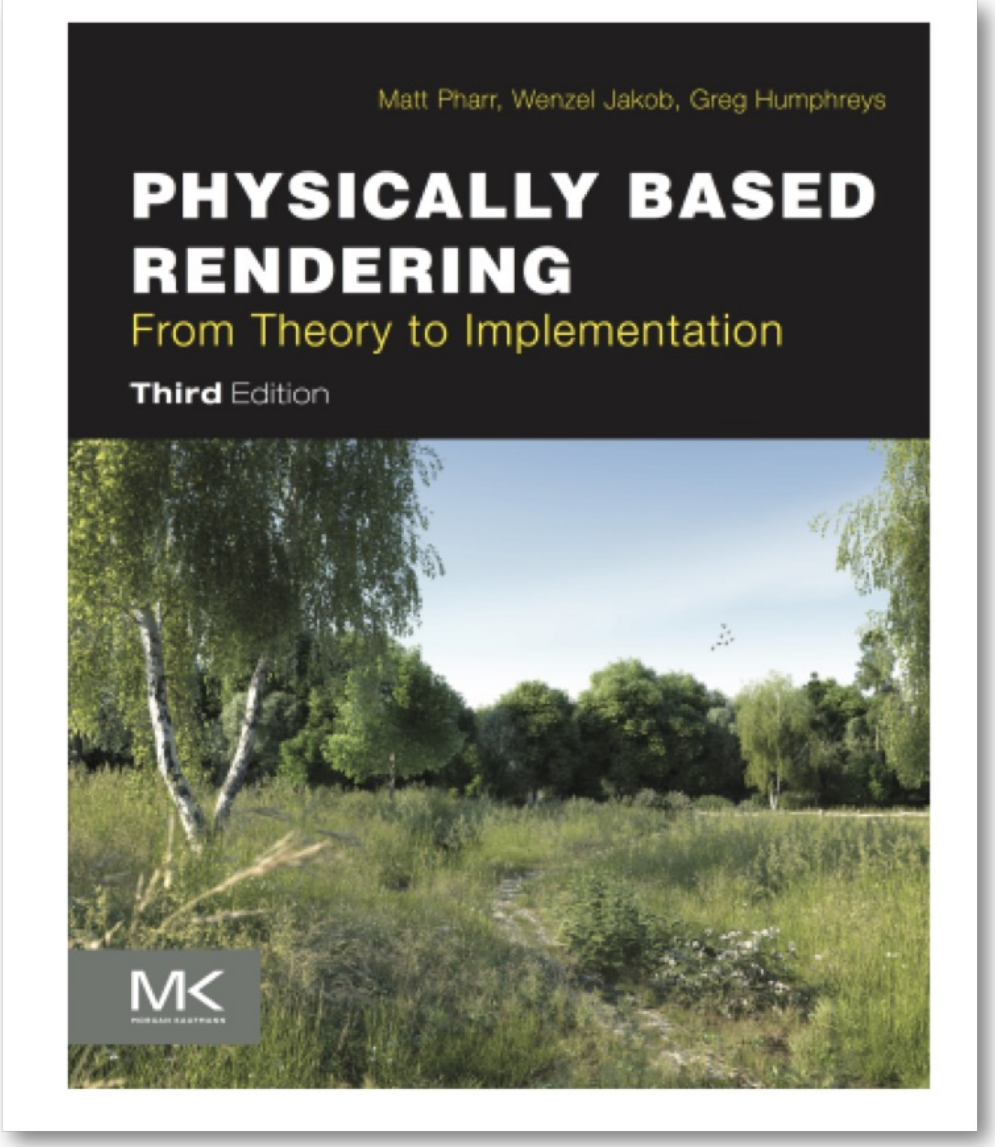
Note: Carla, Anyverse, Omniverse

Quantitative computer graphics is necessary component for materials and lights

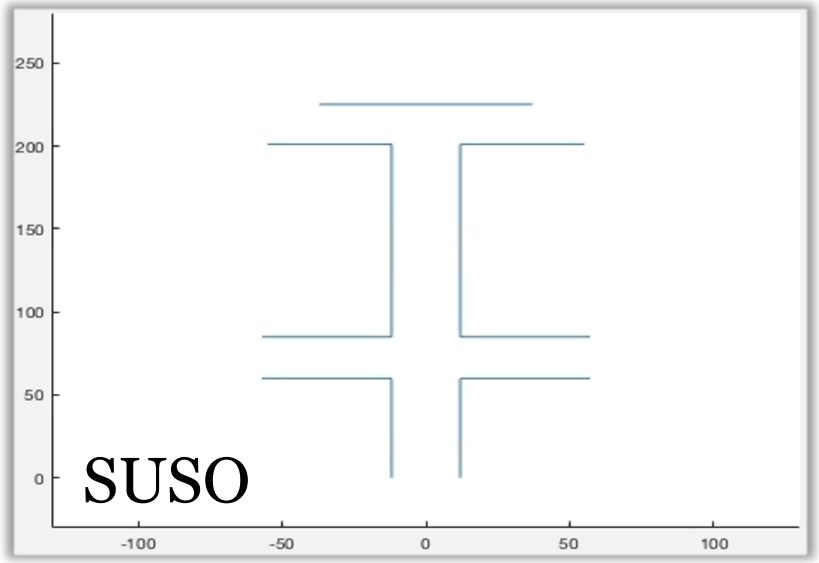
PBRT uses ray tracing from the sensor, through multi-element optics, into the scene spectral radiance. It includes accurate physics and the option to specify physical units

We added methods for

- Diffraction approximation
- Human eye model
- Aspherical lenses
- Microlens array at the sensor (light field)
- Linear models of texture maps to control surface spectral reflectance
- Fluorescence (Medical imaging)
- Participating media (Underwater)



Scene generation is essential



- Traffic flow (Position, orientation, speed of cars, trucks, pedestrians) is **randomly** generated with SUMO by given **parameters.**
- Appearance of the vehicles are **randomly** according to car color popularity.
- Buildings and other static objects are are **randomly** sampled from the database (Flywheel) and designed to place along side of the road reasonably(SUSO) by given **parameters**.
- Camera and lighting parameters are established by given **parameters**

SUMO is an open source, highly portable, microscopic and continuous road traffic simulation

SUMO

ISETCLOUD

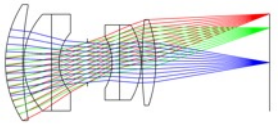
ISET3D

PBRT (docker)

(q/s/sr/nm/m²)

Sensor irradiance

(q/s/nm/m²)



Pixel level depth and labels

Asset management and curation for scene generation is essential

Assets:

Recipe.json:

Contains material, geometry, texture pointer information.

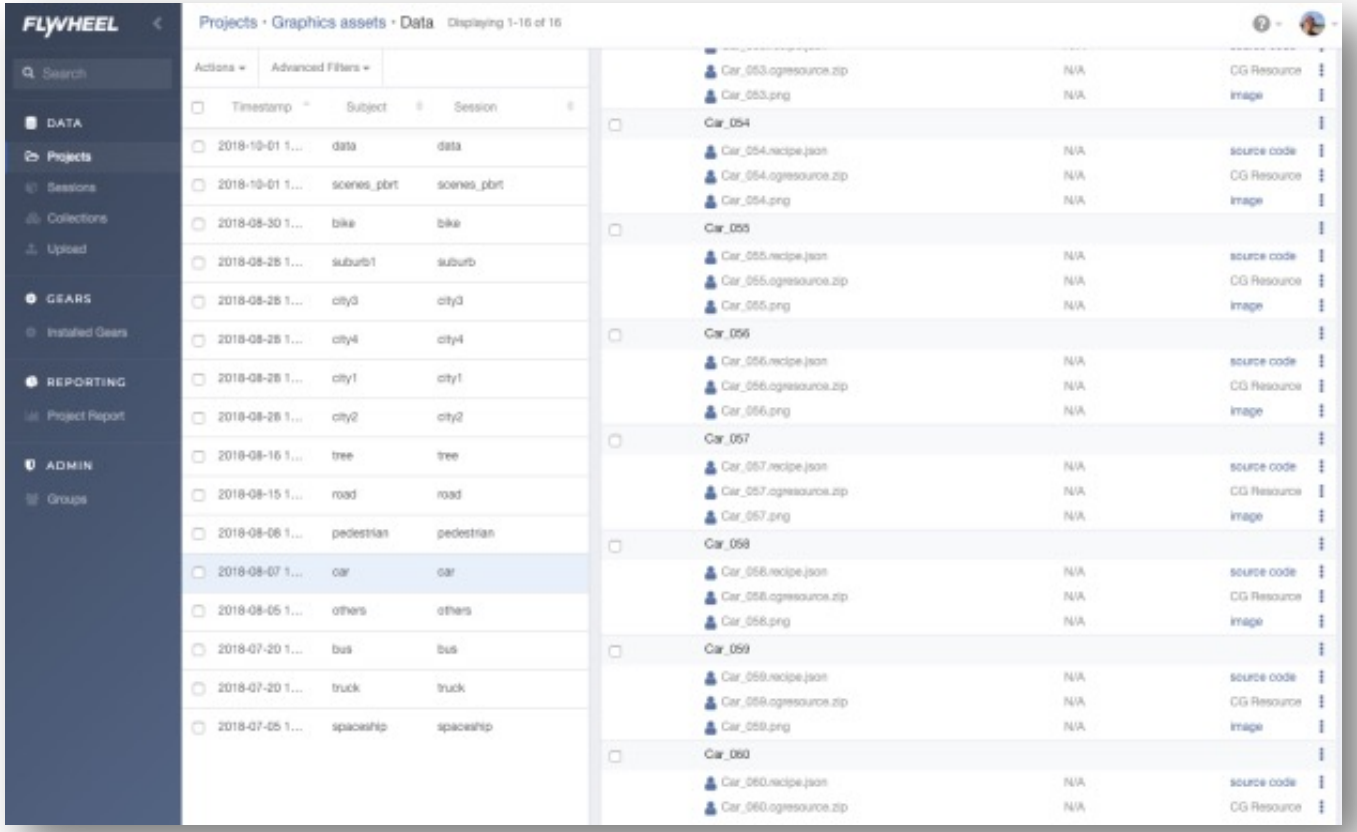
CgResource:

Brdfs of material, texture files, geometry(point, triangle vertices, faces).

Png File:

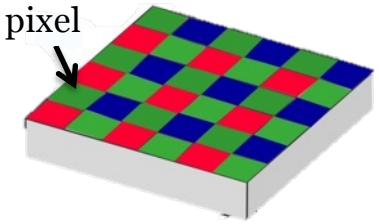
For preview.

Flywheel cloud database with SDK



Imaging sensors will be re-designed for machine perception

Sensors for evaluation



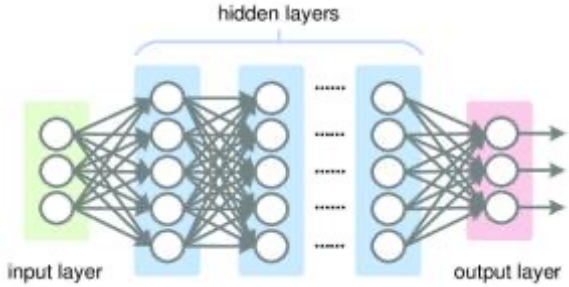
Sensor A

Parameter	Value
Optical Format	1/3 inch
Active Imager Size	4.51 mm(H) x 2.88 mm(V)
Active Pixels	752H x 480 V
Pixel Size	6 um x 6 um
Color Filter Array	RGGB Pattern
Full Resolution	752 x 480
Frame Rate	60fps
Dark Noise	1.0mV/pixel/second
Read Noise	1.0 mV

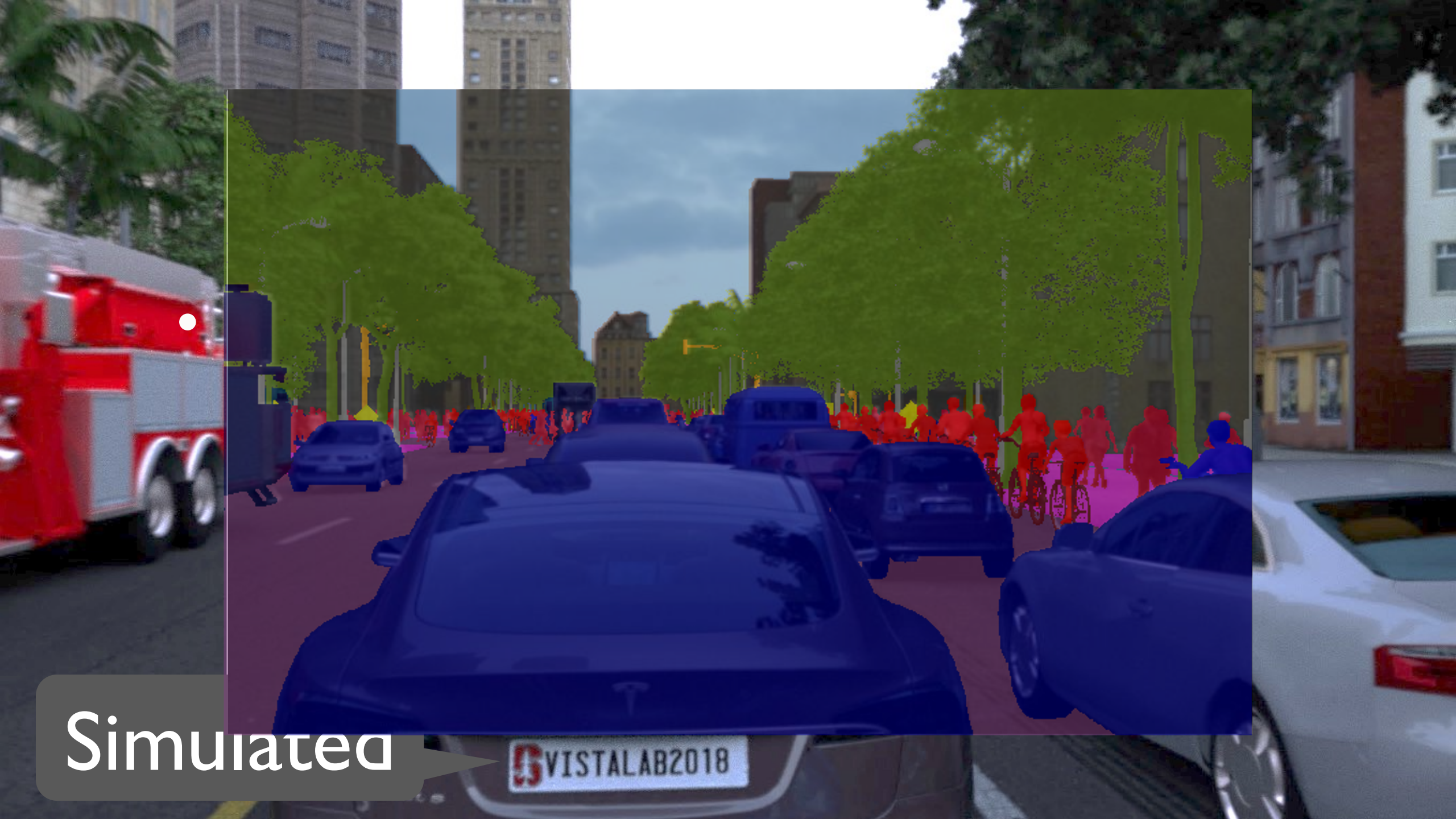
Sensor B

Parameter	Value
Optical Format	1/3 inch
Active Imager Size	4.51 mm(H) x 2.88 mm(V)
Active Pixels	1504H x 960 V
Pixel Size	3 um x 3 um
Color Filter Array	RGGB Pattern
Full Resolution	1504 x 960
Frame Rate	60fps
Dark Noise	1.0mV/pixel/second
Read Noise	1.0 mV

CNN models

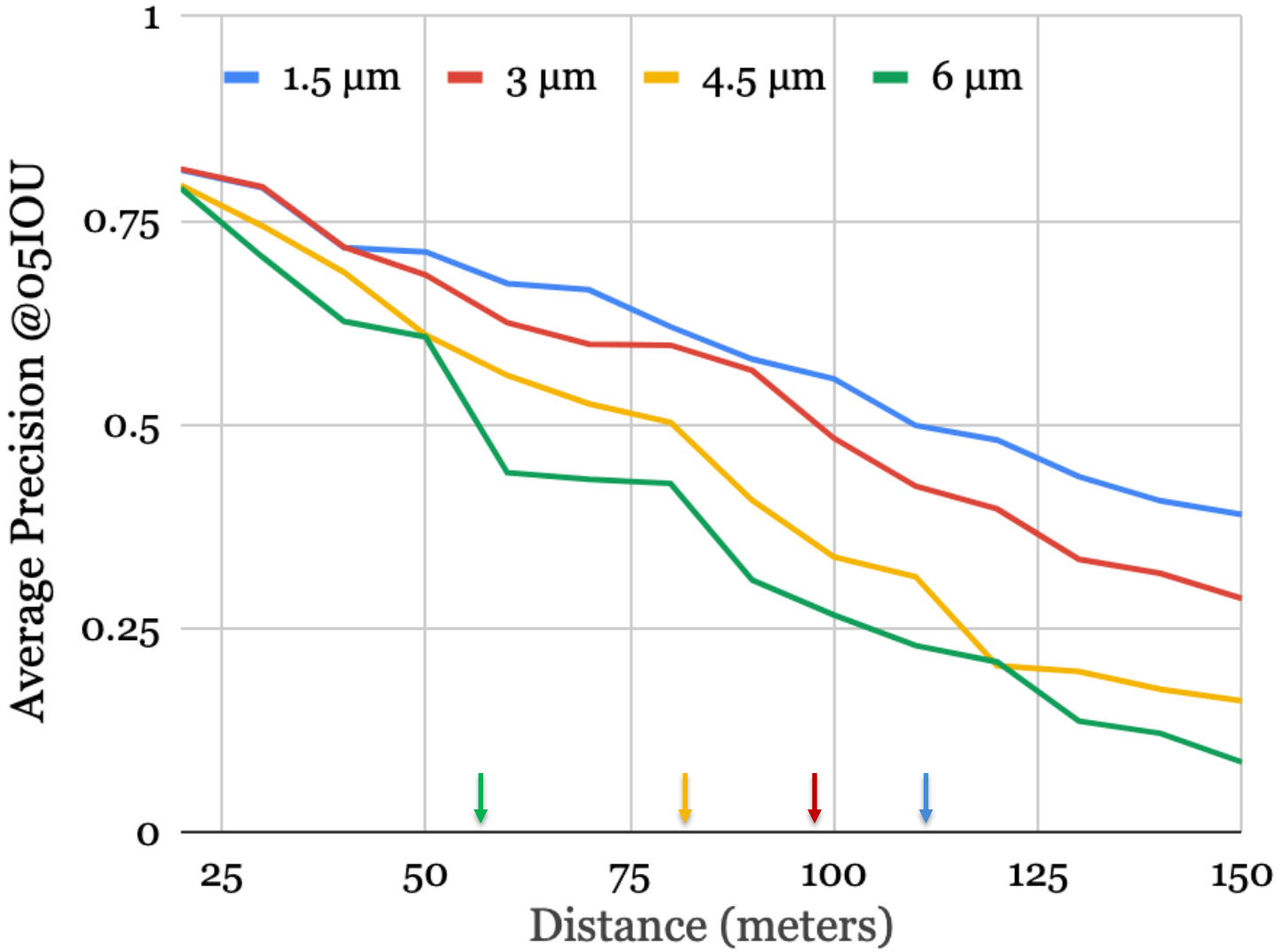


Parameter	Value
Feature Extractor	faster_rcnn_resnet101
Num of Layers	101
First stage features stride	16
First stage nms iou threshold	0.7
First stage max proposals	300
Optimizer	momentum_optimizer
Regularizer	l2_regularizer
Score converter	SOFTMAX

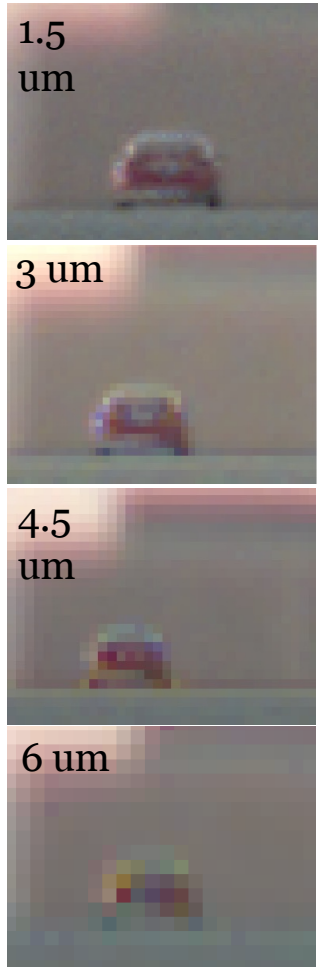

VISTALAB2018
Simulated

Meaningful performance metrics computed from simulation

- This analysis simulates average precision of a CNN (ResNet-50) comparing with cameras with different pixel sizes
- The accuracy depends on pixel size, systematically
- The simulation includes meaningful units
- The performance of the 1.5um pixels is similar to current LIDAR detectors

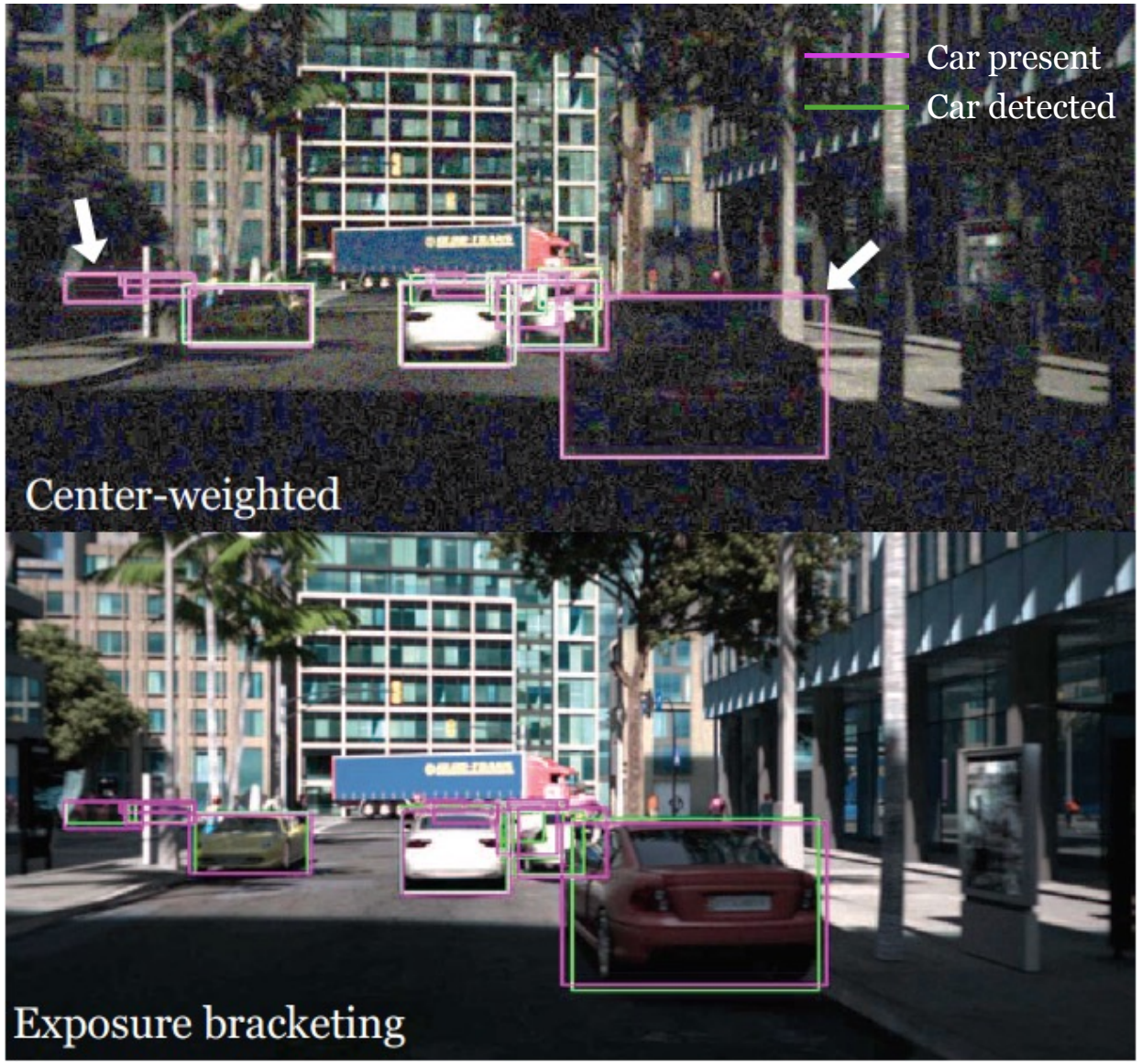


Car at 150 m



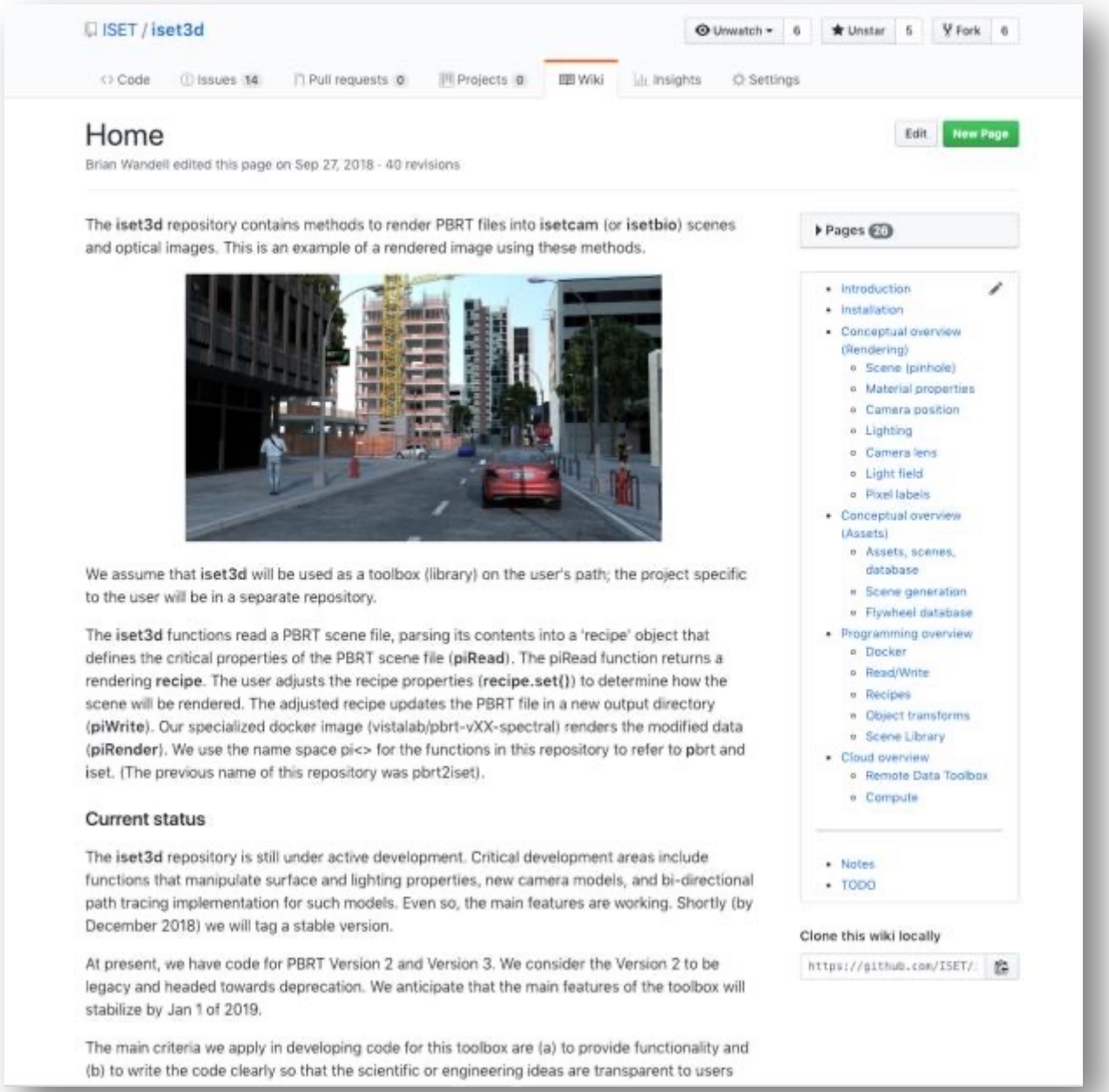
Examining edge cases: exposure algorithms

- Another value of simulation is the ability to look for edge cases
- In our work we have compared different exposure algorithms: center-weighted and exposure-bracketing are shown here
- In this case the center-weighted failed to see a car that was detected by exposure bracketing
- Examining such cases helps us understand the weaknesses of different approaches



Simulation for camera design in automotive machine-learning

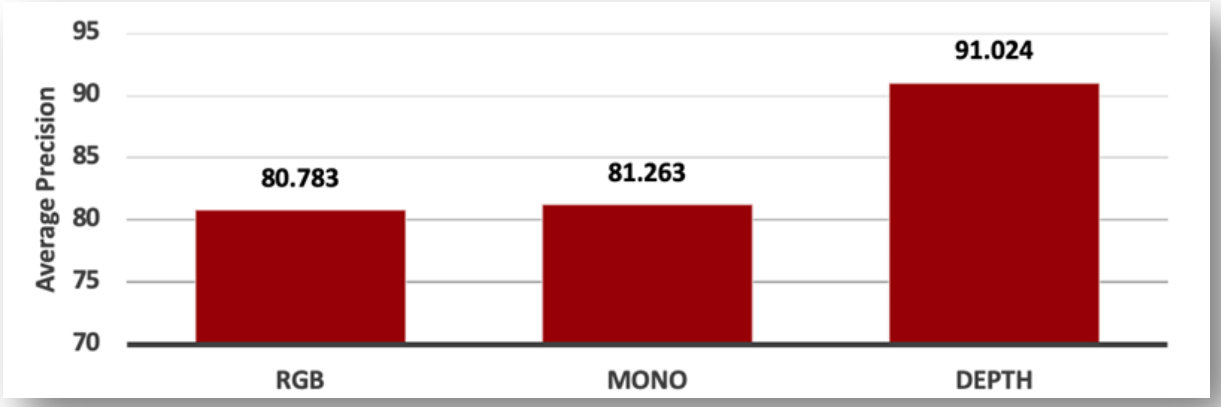
- A purely empirical approach to camera design for automotive application is impractical.
- We created an end-to-end system to automate the generation of physically accurate natural scenes and sensor images for network training and evaluation
- We are experimenting with different camera designs



Next: Depth, LIDAR, Time-of-Flight

- After equating for spatial resolution, depth information is at least as valuable as radiance information for scene segmentation.

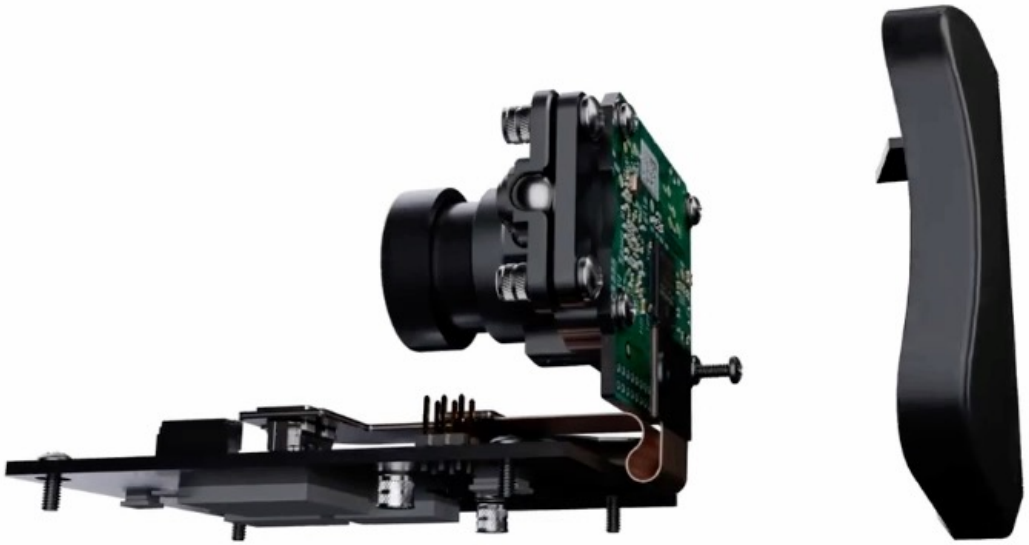
- Combining depth and radiance by inserting the depth map into an image input channel increases the average precision of vehicle detection substantially.



ISETAuto: Detecting vehicles with depth and radiance information

ZHENYI LIU[1], JOYCE FARRELL[2] AND BRIAN WANDELL[2]
[1]State Key Laboratory of Automotive Simulation and Control, Jilin University (e-mail: zhenyiliu27@gmail.com)
[2]Stanford University (e-mail: jefarrel, wandell@stanford.edu)
Corresponding author: Zhenyi Liu (e-mail: zhenyiliu27@gmail.com)

Brightway vision hardware to evaluate and simulate



VISDOM.
ALL-WEATHER AUTOMOTIVE CAMERA SYSTEM

Gated-CMOS time-of-flight camera with NIR illuminator

Nighttime and fog

Summary point

Quantitative computer graphics involves skills that are beyond the training in most vision science labs and early commercial ventures (startups).

The community can benefit from open-source, physically accurate, graphics tools that we trust and understand.

Simulation technologies for image systems engineering

Brian A. Wandell
Stanford Center for Image Systems Engineering (SCIEN)

Wu Tsai Neurosciences Institute (WTNI)
Stanford Center for Cognitive and Neurobiological Imaging (CNI)