

Simulation technologies for image systems engineering and biology

Brian A. Wandell

Stanford Center for Image Systems Engineering

Wu Tsai Neurosciences Institute
Stanford Center for Cognitive and
Neurobiological Imaging

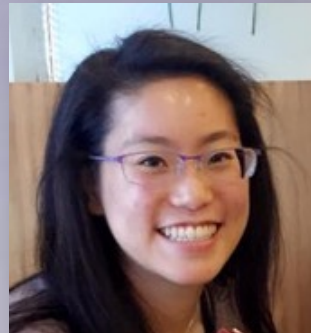
QUANTITATIVE MEASUREMENTS

∞

COMPUTATIONAL MODELS

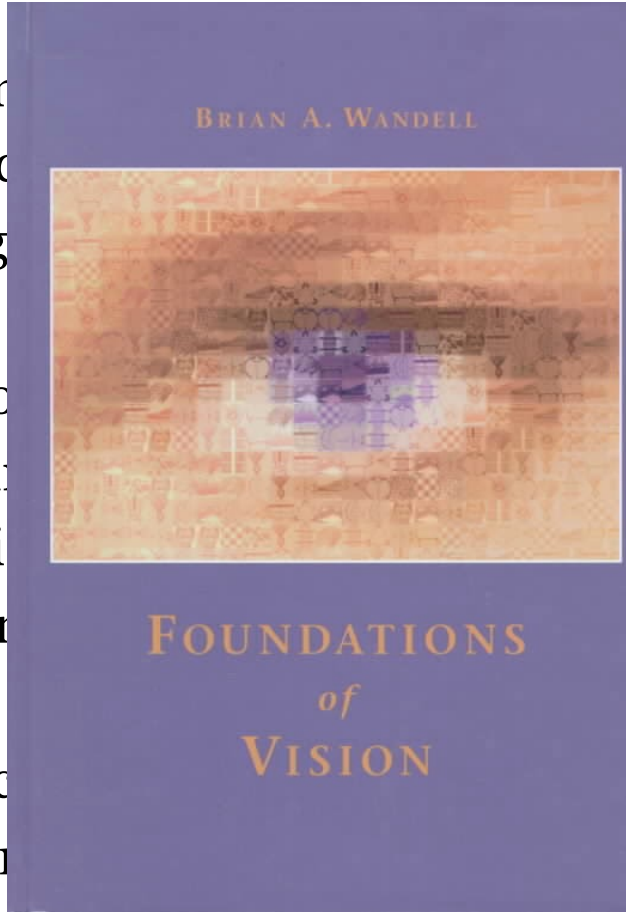
∞

CHECK AND SHARE



Vision science

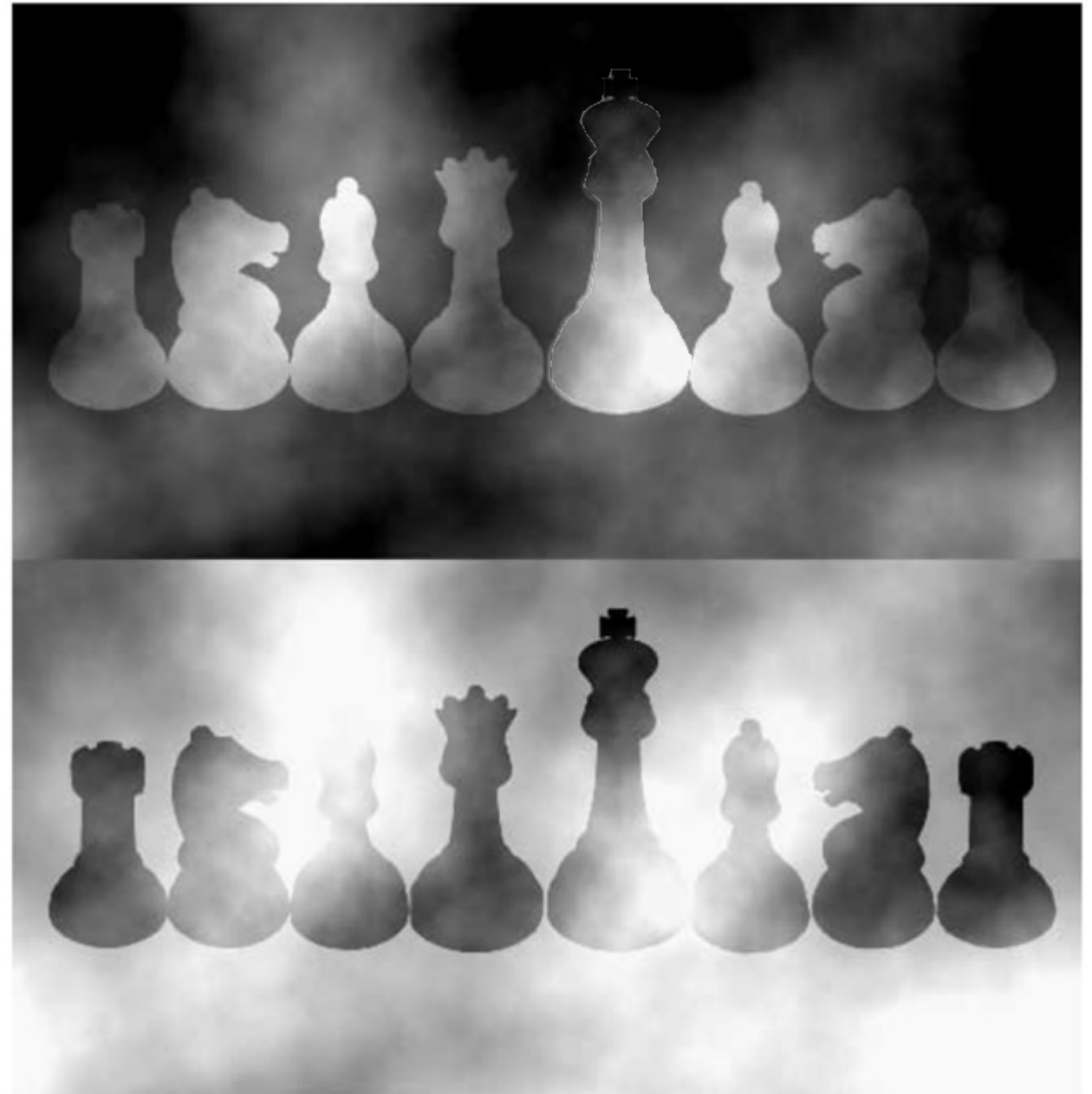
- Even simple tasks - depend on the image
- The visual system develops through neuroscientific intelligence
- Vision science is important for the environment



brightness
ion of
circuits

tial in

portant



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!

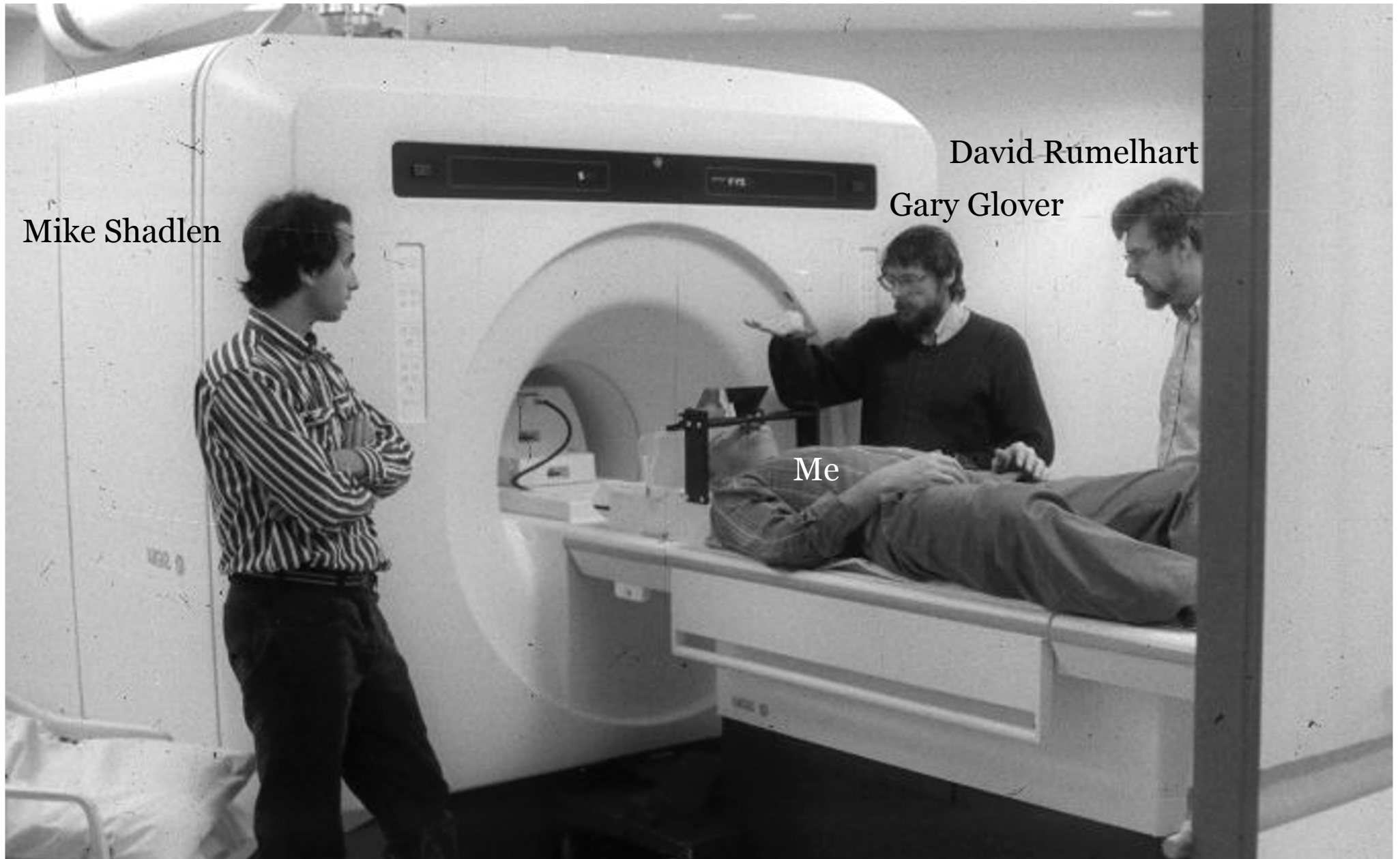


- Neurons/mm³: 10^4 - 10^6
- Cortical Neurons: 10^{11}
- Synapses/neuron: 10^3
- Cortical Synapses: 10^{14}
- Surface area of each hemisphere: 20 x 30 cm²

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

Image from Graham Johnson

VISTA Vision Science: MRI – September, 1991



Mike Shadlen

David Rumelhart

Gary Glover

Me

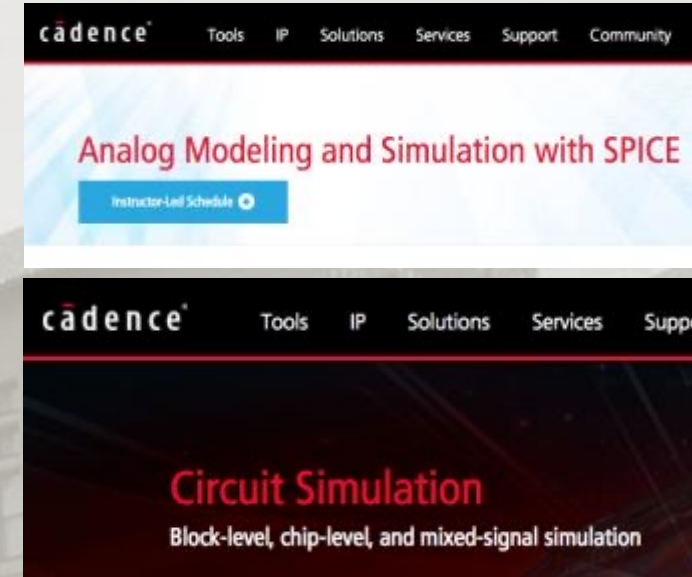
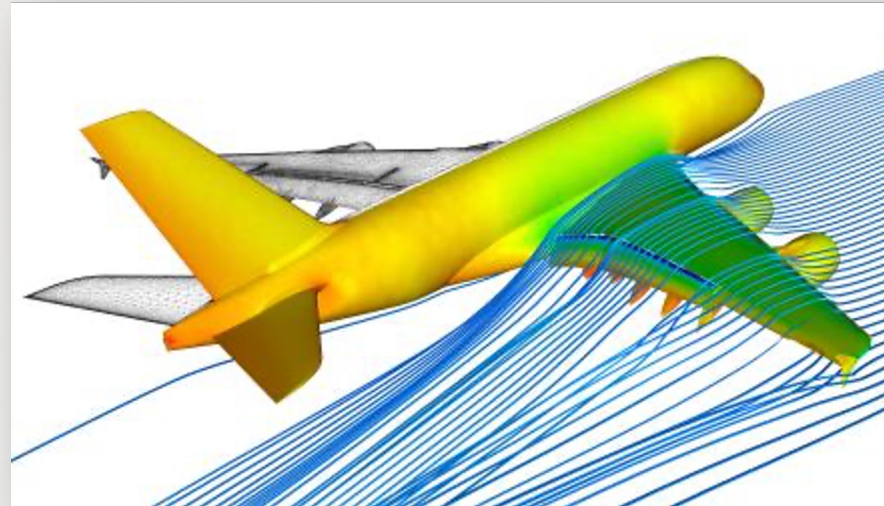
Founding director of the MRI center across the street in Building 420

<https://vimeo.com/705176118/c6a9850b87>



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

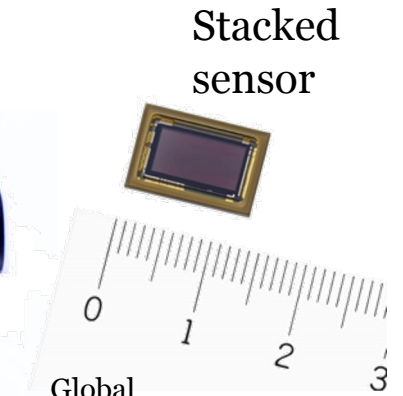
Integrated circuitry

Simulation is particularly useful for industry innovations

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



Multiple lens



Stacked sensor



RGB-depth



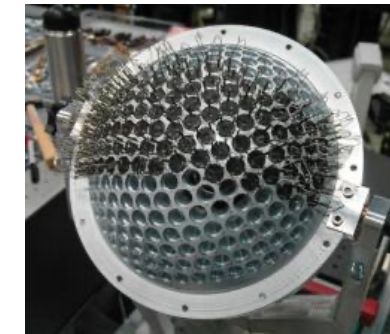
Global shutter



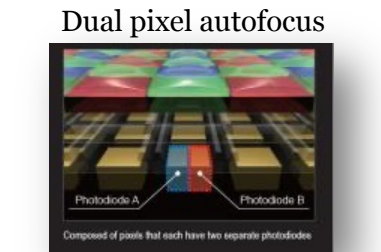
Light field



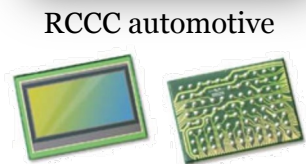
360 Surround Video



Massive resolution



Dual pixel autofocus



RCCC automotive

Standard software exists for simulation of separate components

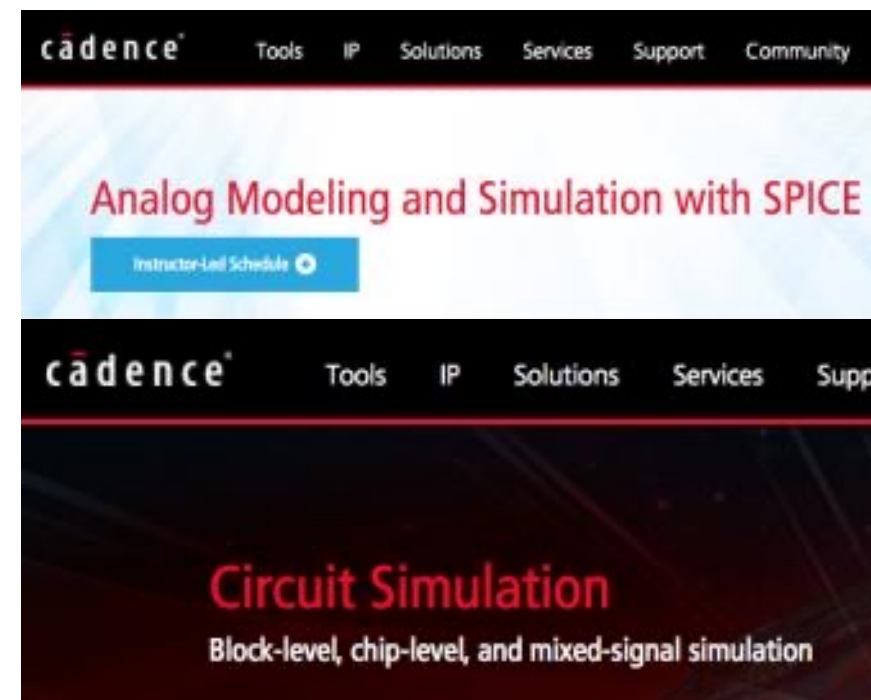
Lens design



Silicon design and Photonic Structures



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

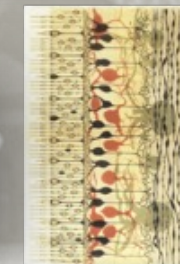
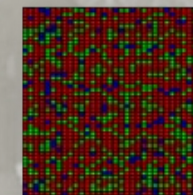
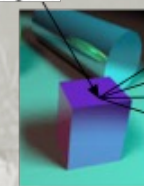
Open Sourced on GitHub
in 2018



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): Starting from 2018

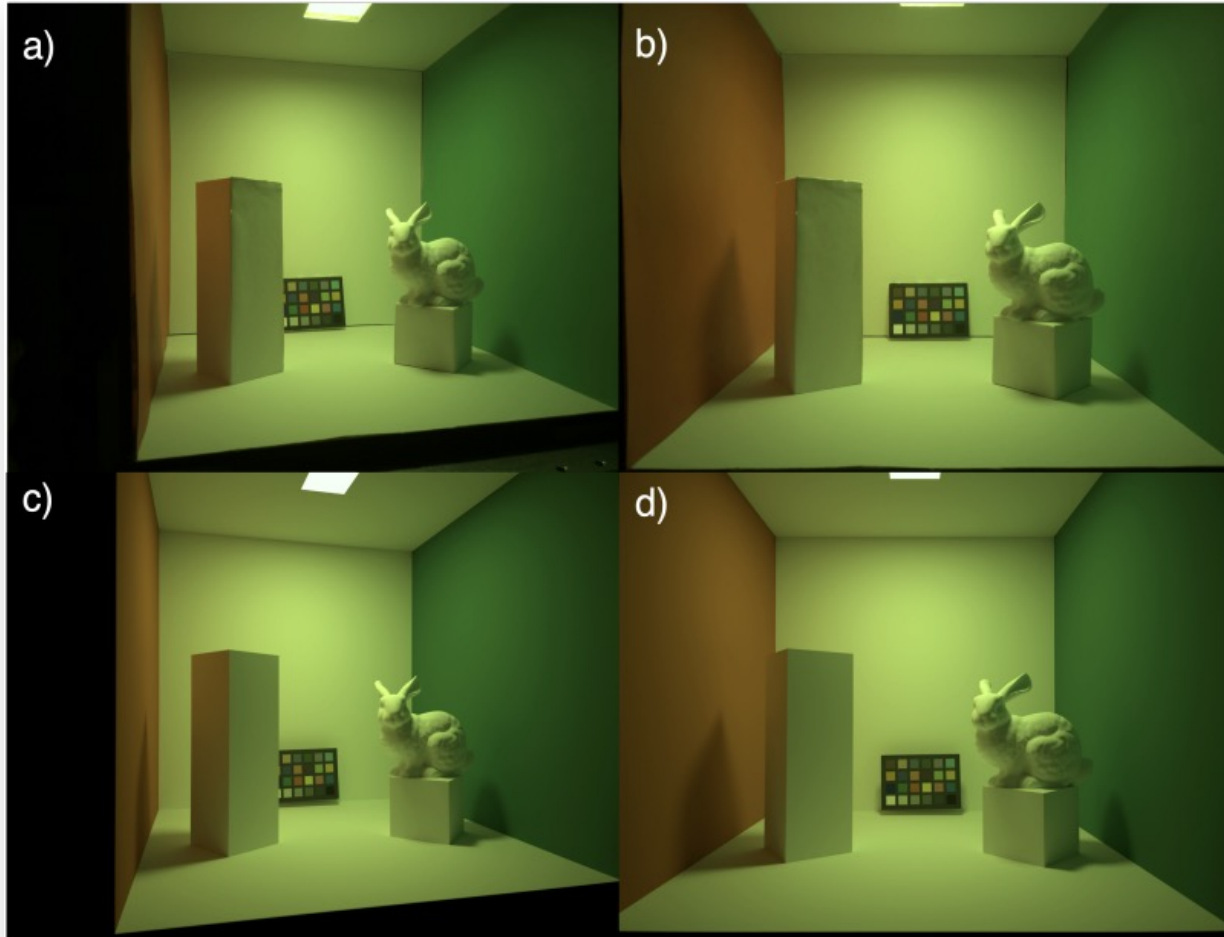
- **ISET3d** is a Matlab toolbox that uses **PBRT** to calculate three-dimensional scene spectral radiance and sensor irradiance of complex scenes
- **ISETCam** is a Matlab toolbox that computes the image sensor response from the sensor irradiance; it also includes many industrial tools for evaluation of color (CIE standards) and spatial resolution (ISO standards)
- **ISETBio** is a Matlab toolbox that computes the visual encoding



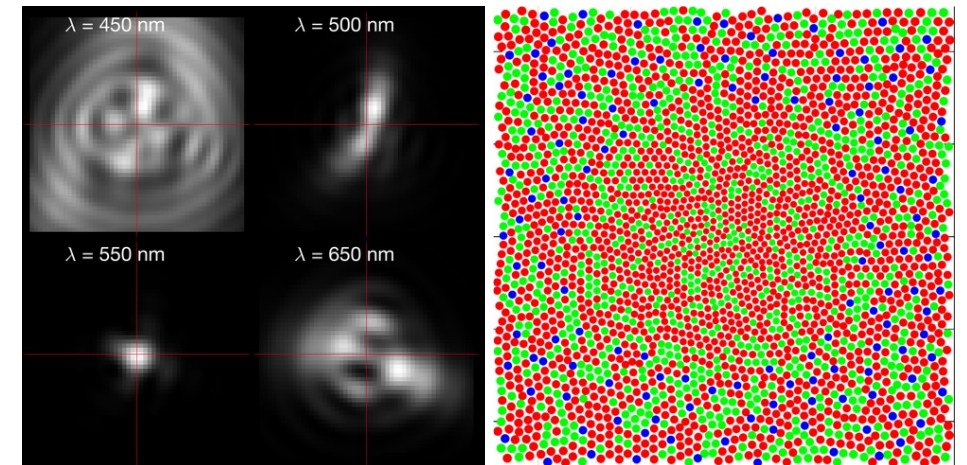
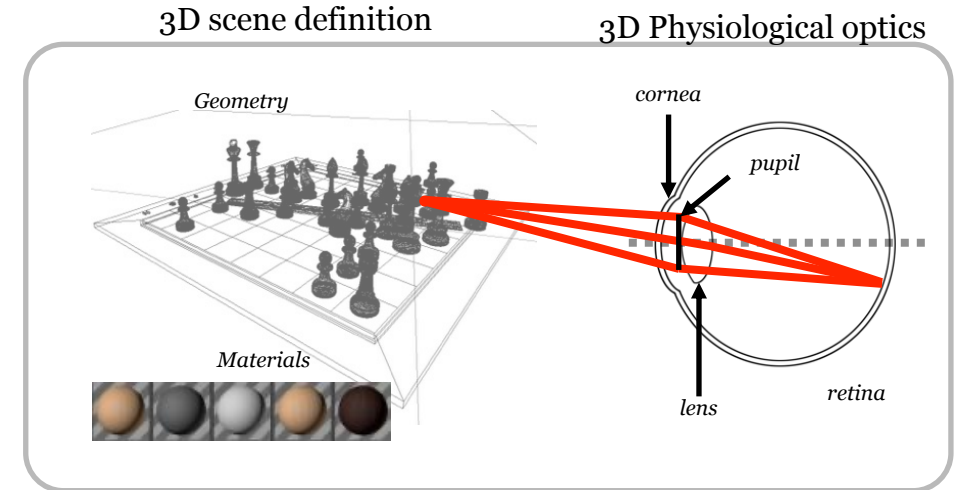
Presentation outline

Part 1: Validation and applications

Two viewpoints: Simulated and measured



Part 2: Visual encoding



What goes into a 3D image systems simulation

More about these components during the presentation

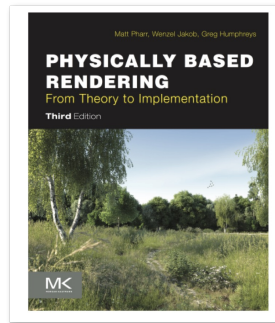
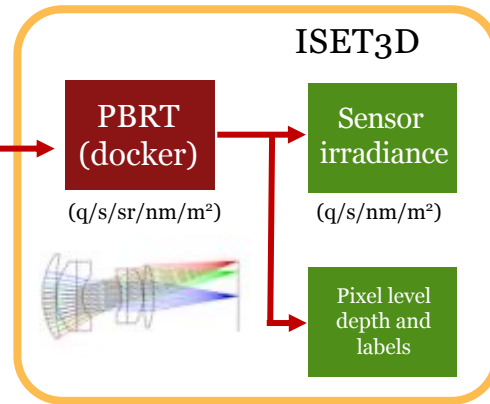
Asset generation



Scene Auto-assembly



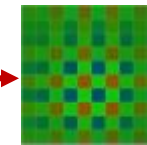
Scene spectral irradiance rendering



Sensor modeling

ISETCam

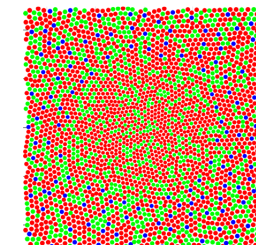
(electrons/pixel)



Pixel-level labeled sensor data

Annotation

ISETBio

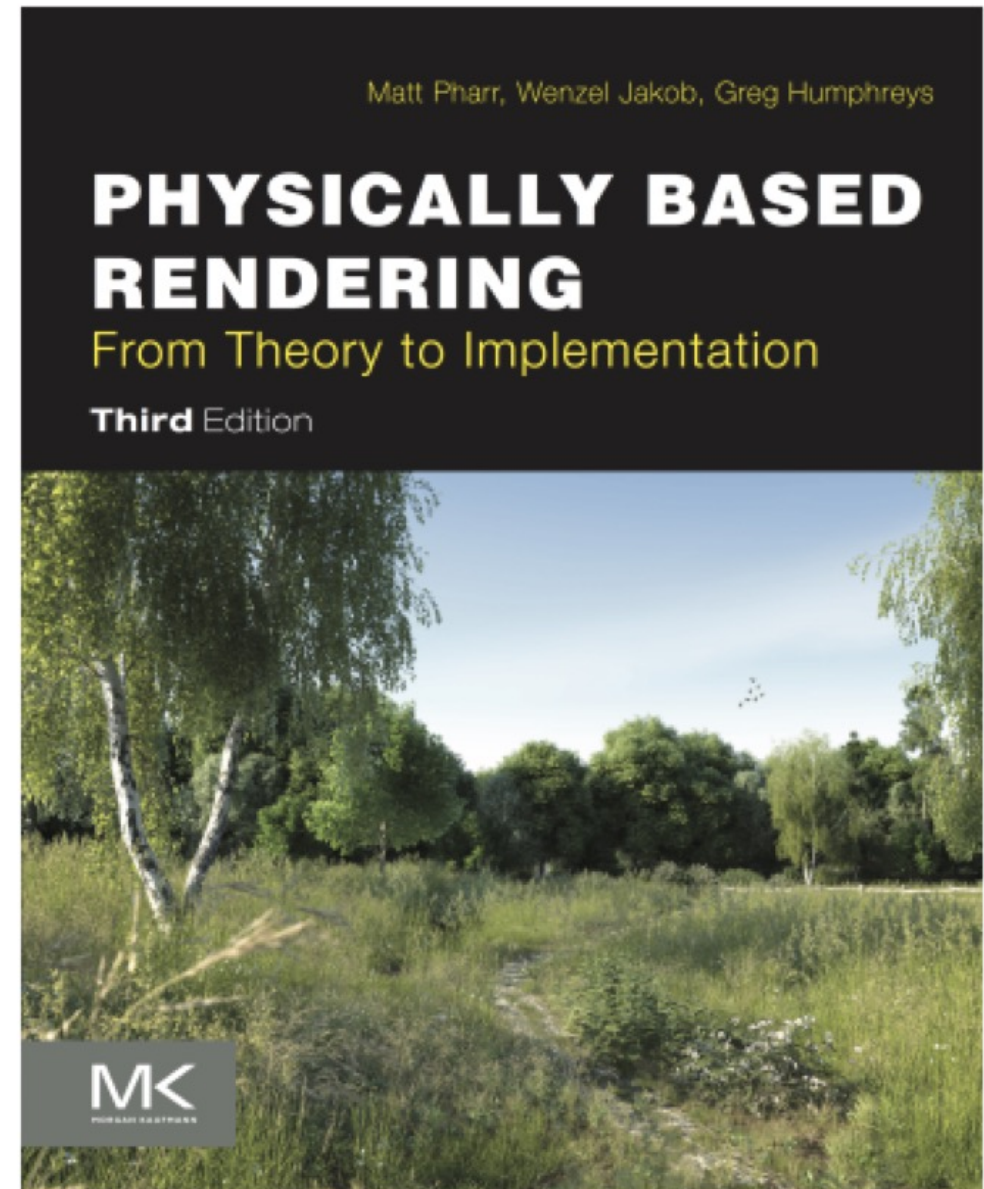


Training

PyTorch

TensorFlow

- Progress in computer graphics enables us to create synthetic and yet highly realistic input data.
- We want simulations with meaningful units; quantitative computer graphics
- Open-source and well documented!

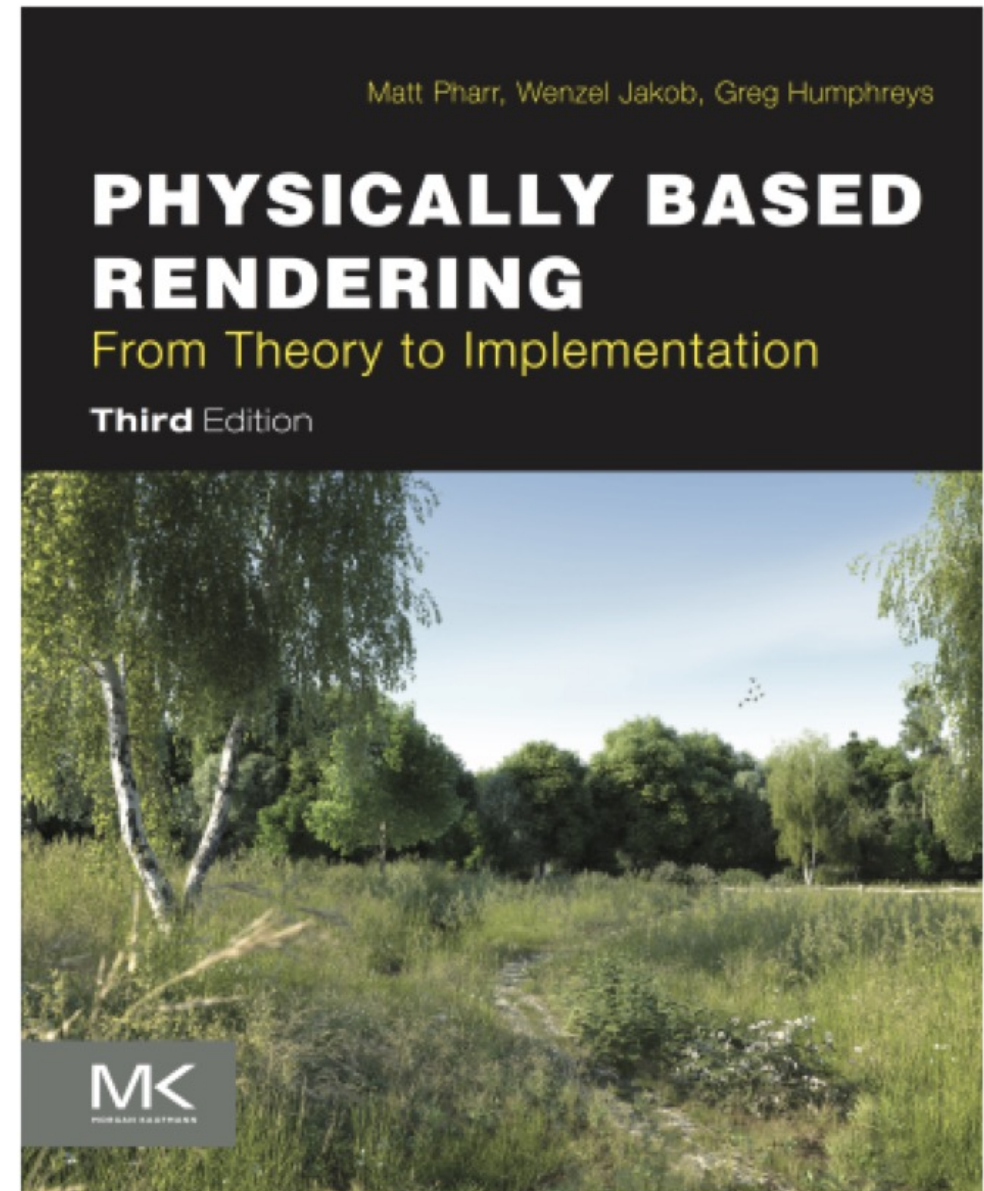


Quantitative computer graphics is necessary component for materials and lights

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 to model and compute

- Diffraction
- Human eye
- Aspherical lenses
- Microlens arrays
- Linear models of texture maps to control surface spectral reflectance
- Fluorescence (Medical imaging)
- Participating media (Underwater)
- Computational imaging (CNN, Ideal observer)



Rasterization is excellent for many purposes, but not for this purpose

- No physical quantities (e.g., spectral radiance, irradiance)
- Pinhole, not real optics
- Bag of tricks for visual appeal

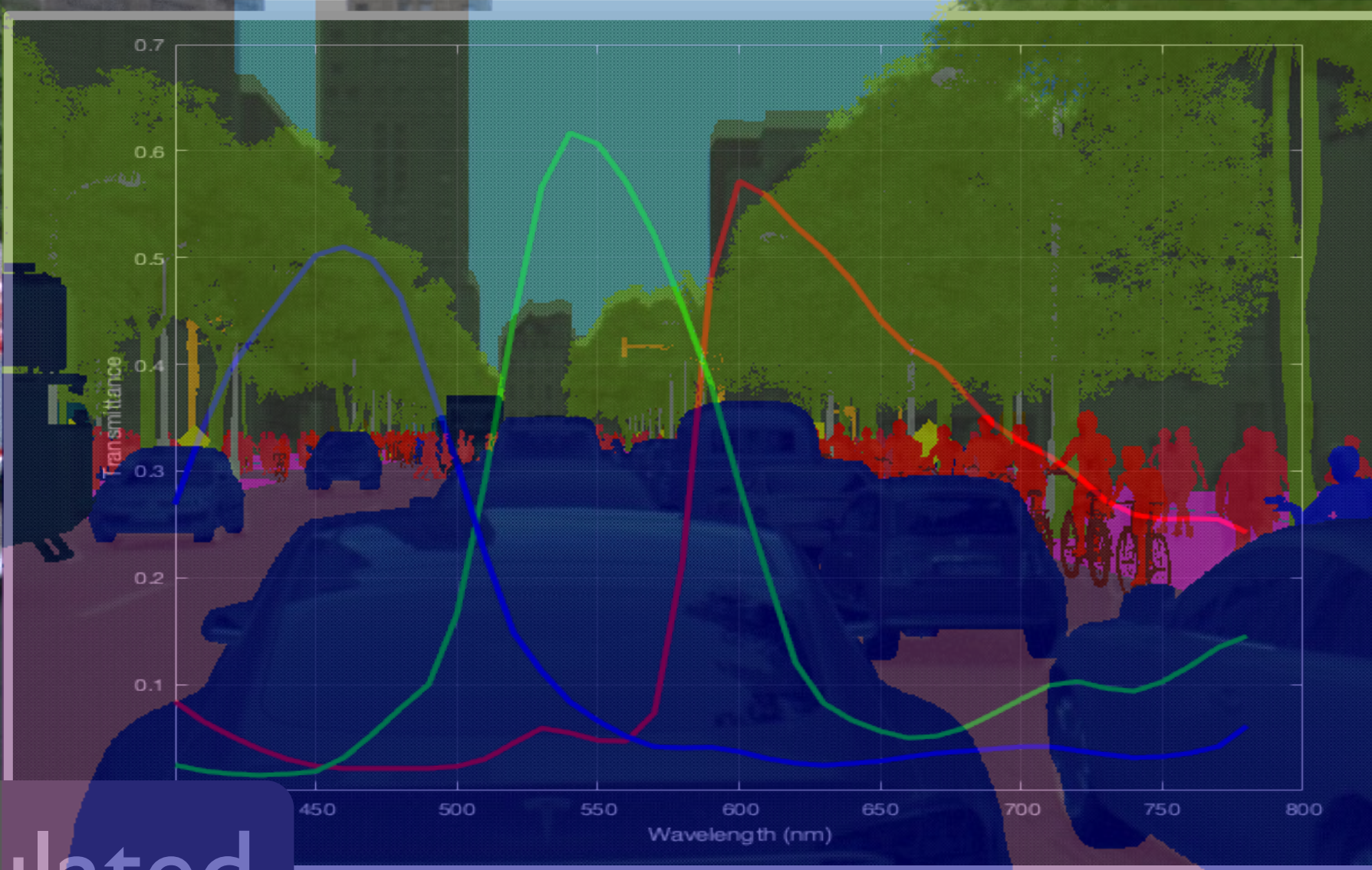
High quality
rasterization –
hand-edited
(800 x 421)
(Unity)



- Attempts to be physically accurate
- Incorporate lens and microlens models
- Produces complex visual effects

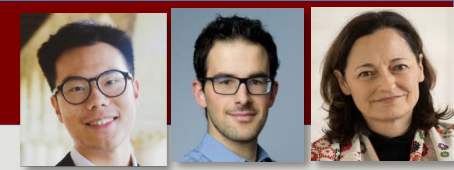
Ray traced –
(712 x 395)
(PBRT)





Simulated

VISTALAB2018

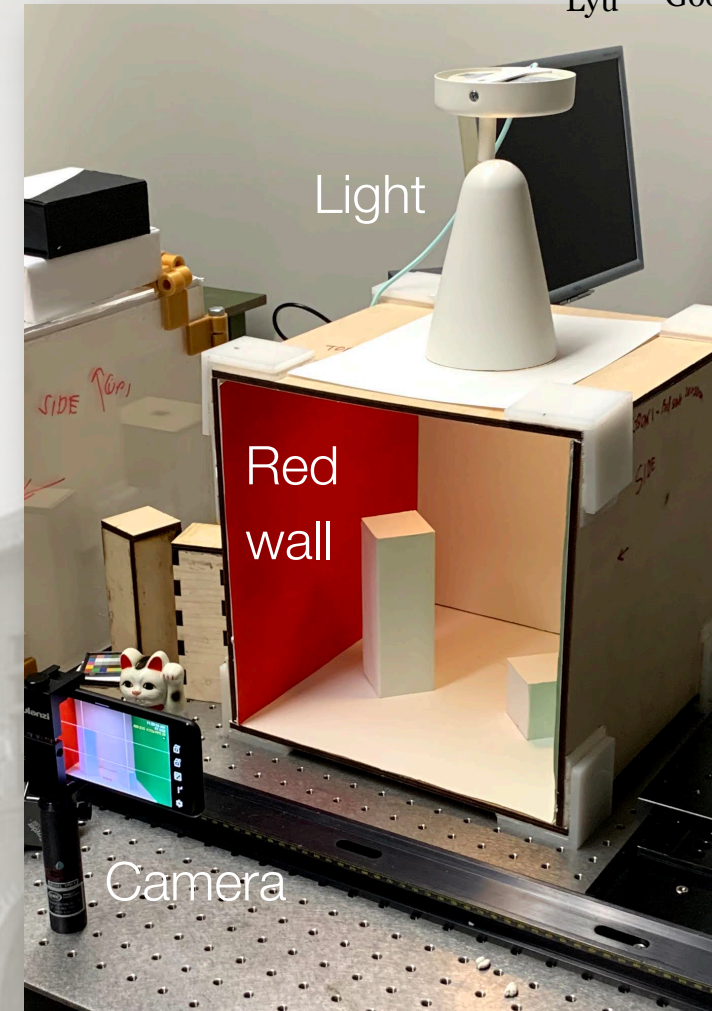


Zheng
Lyu

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



Validation tests: Cornell Box Construction

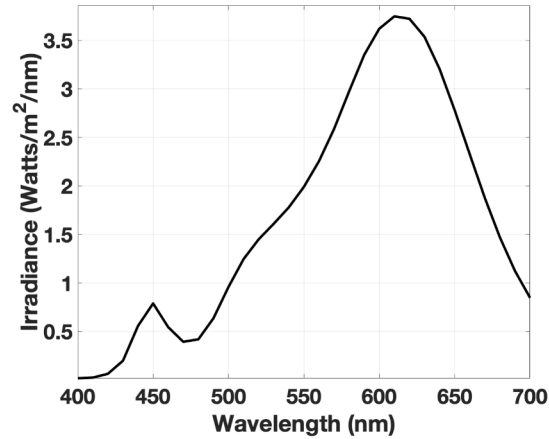
- The Cornell Box was developed to qualitatively test the **accuracy of computer graphics rendering**
- We use it to quantitatively test end-to-end simulation of **3D scenes, optics, and image sensors**



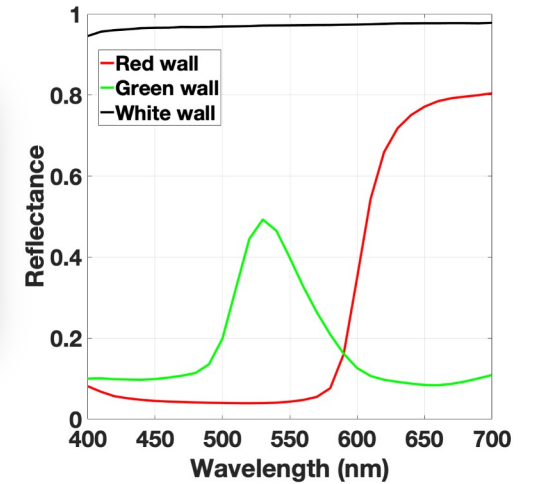
- Goral, Torrance, Greenberg, and Battaile, "Modeling the Interaction of Light Between Diffuse Surfaces," Computer Graphics (Proc. SIGGRAPH 84), Vol. 18, No. 3, July **1984**, pp. 213-222
- Meyer, Rushmeier, Cohen, Greenberg, and Torrance, "An Experimental Evaluation of Computer Graphics Imagery" ACM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

We built one (the 3D scene)

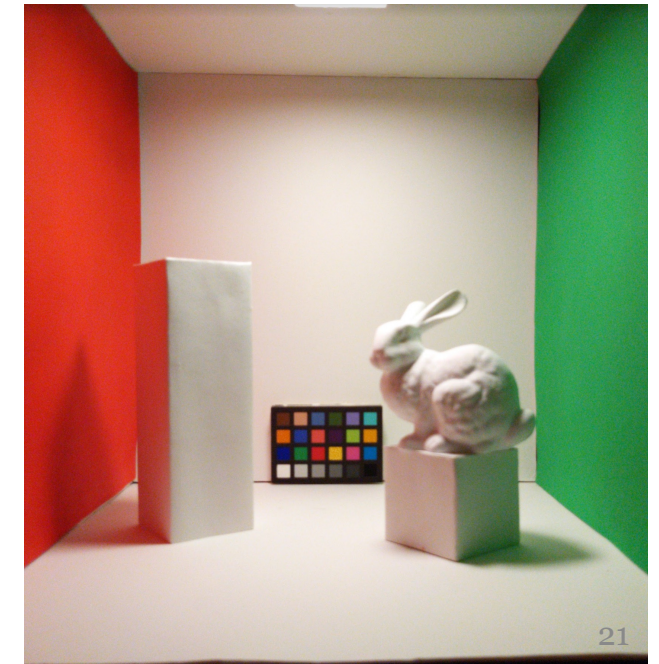
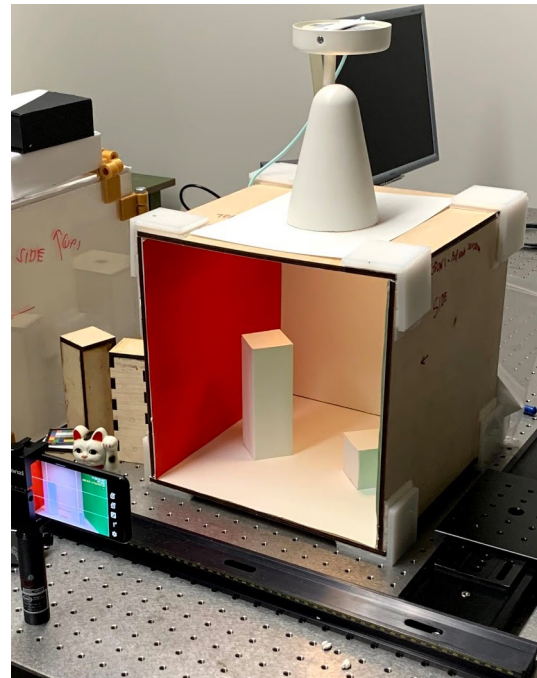
- We measured the asset sizes and scene geometry, surface reflectance, illuminant spectral power distribution



Illuminant SPD

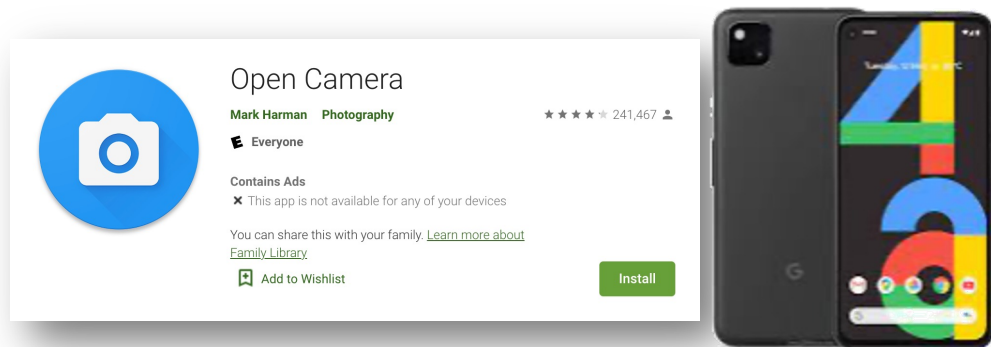


Surface spectral reflectance



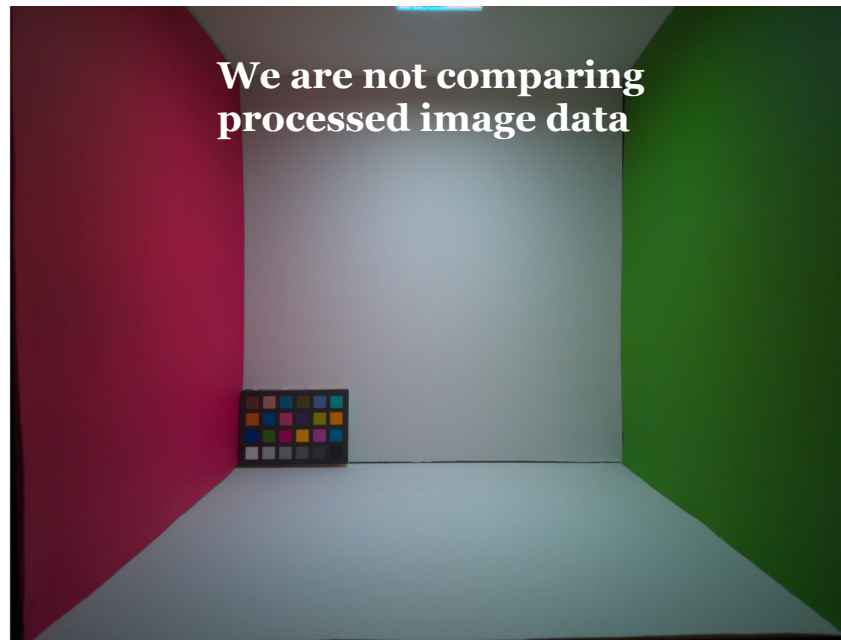
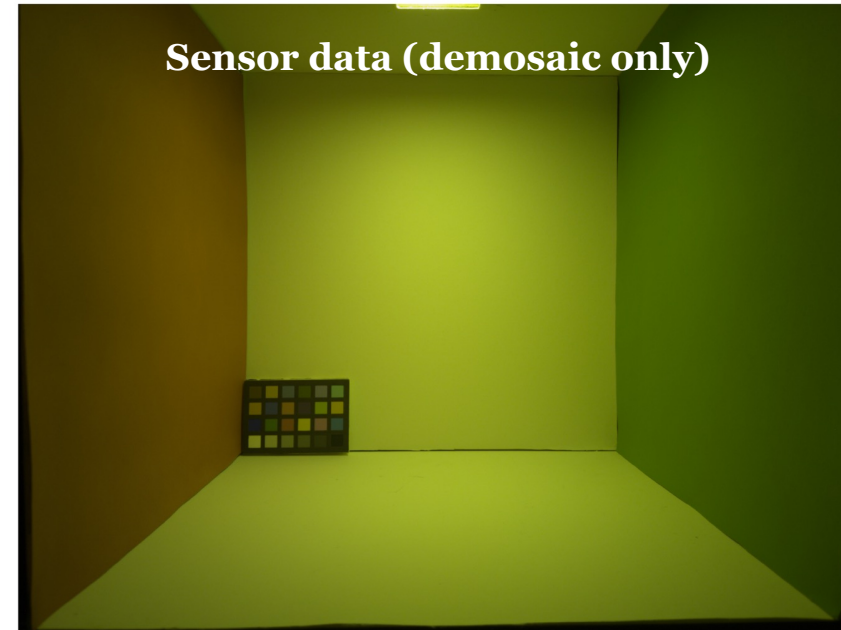
Validation test: Raw sensor data from the camera

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



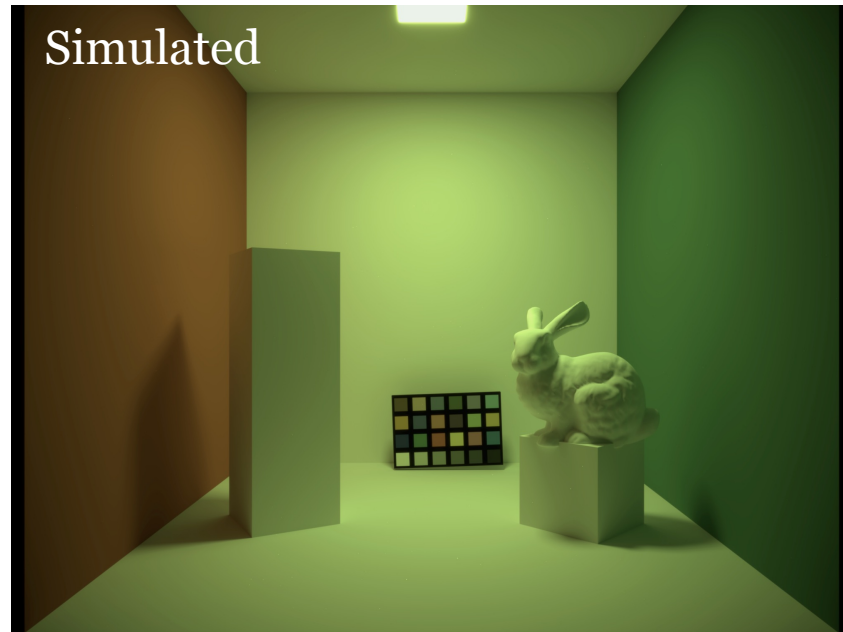
Sensor renderings

- 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

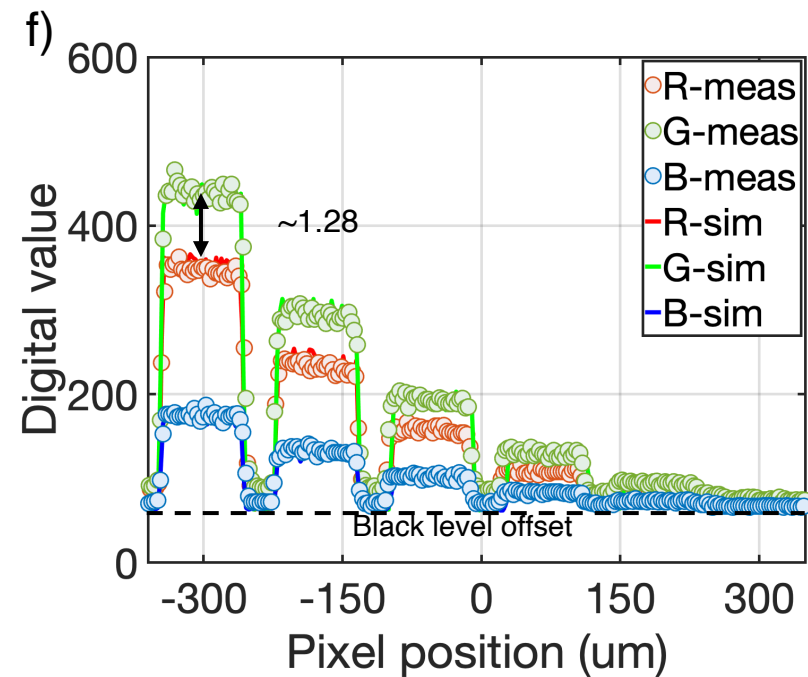
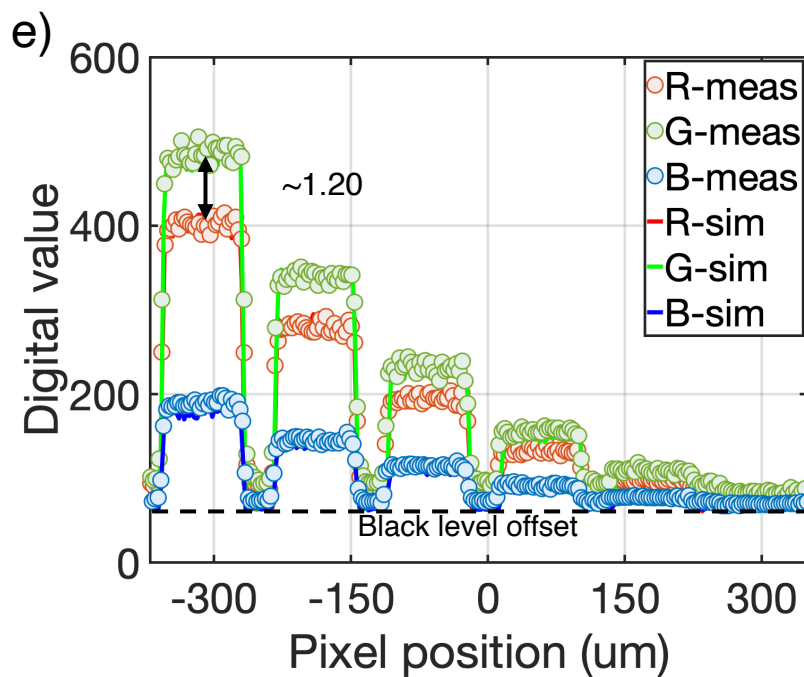
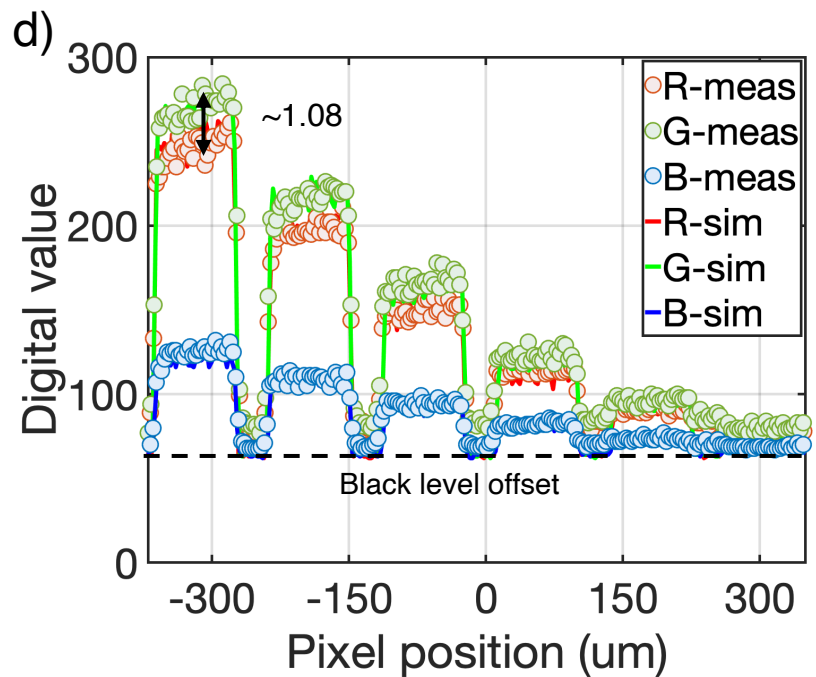
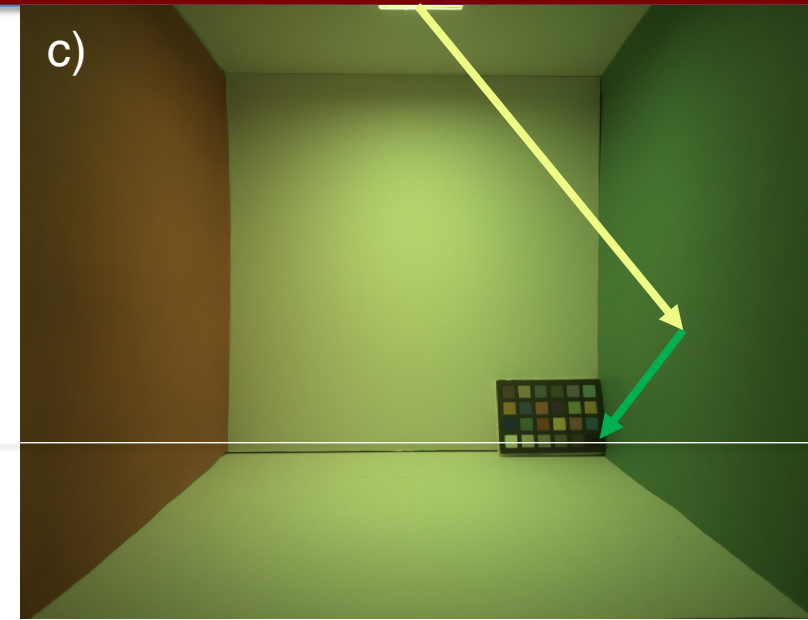
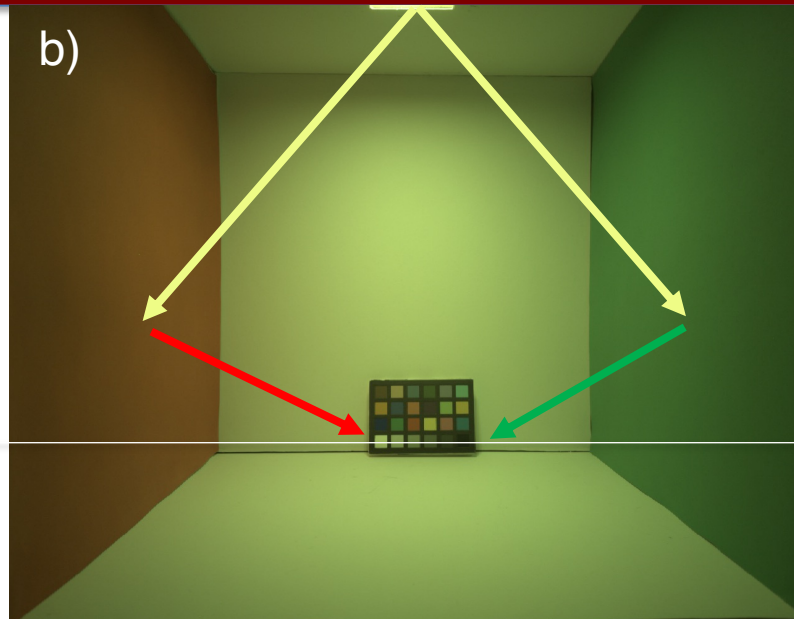
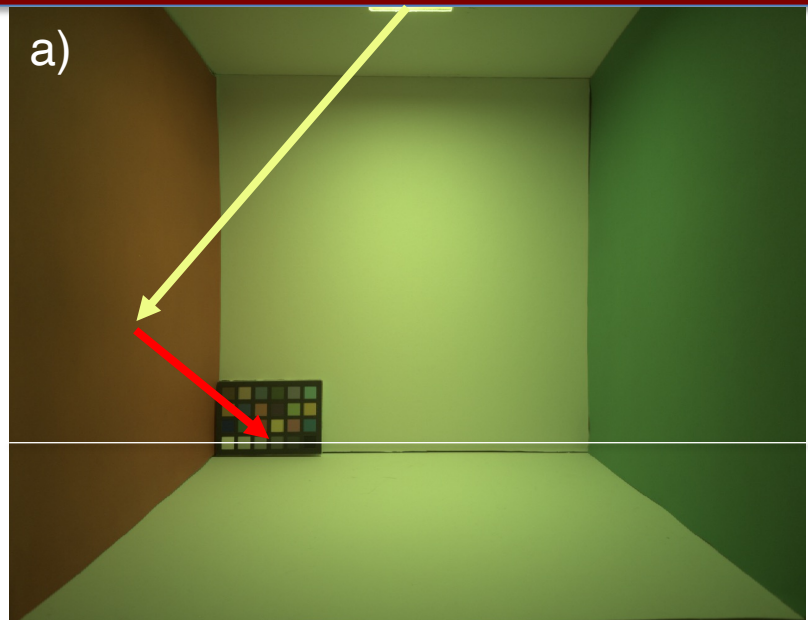


Quick check based just on appearance

- 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!



Validation of the 3d scene representation: Surface inter-reflections

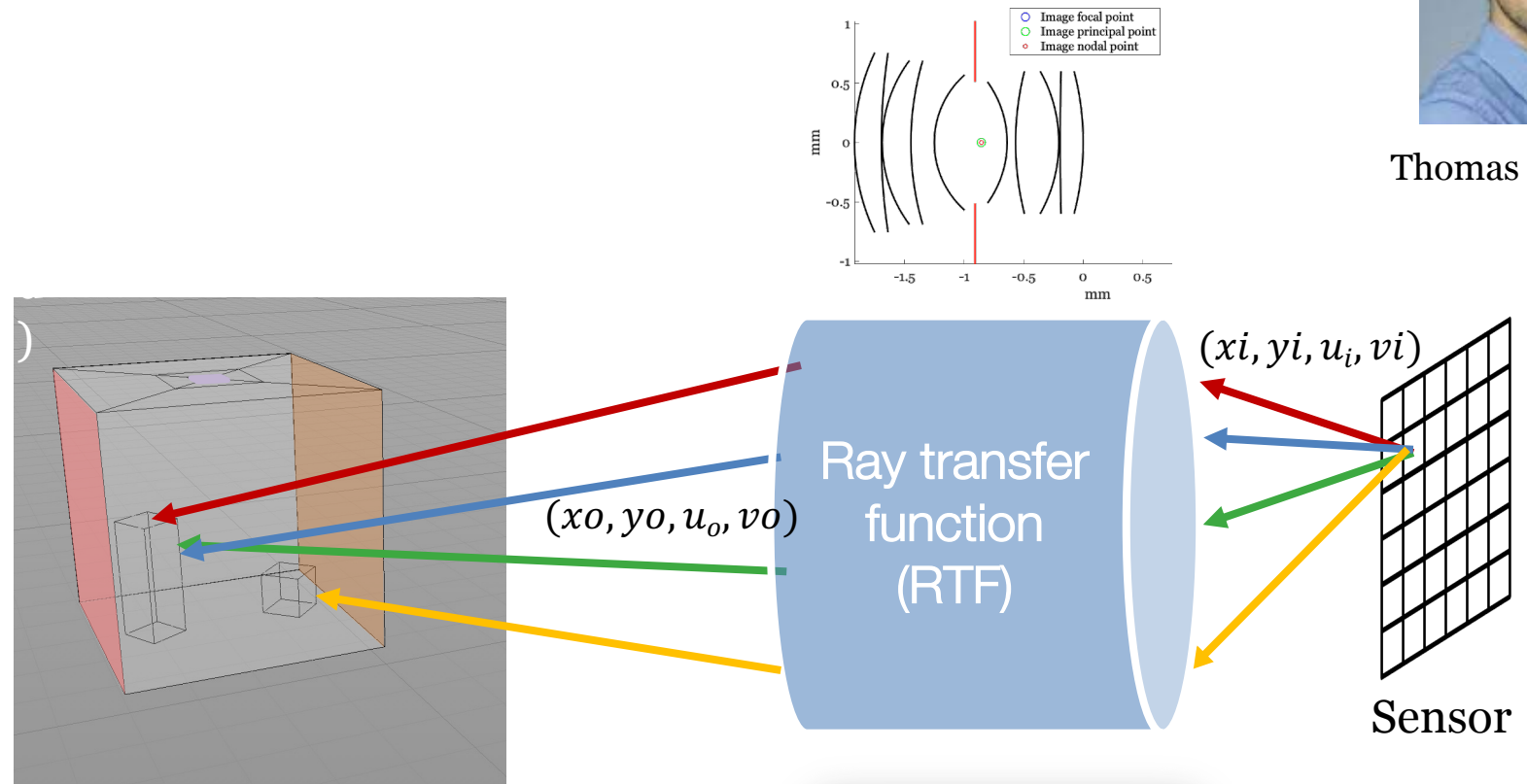


The problem of modeling optics: Ray transfer function



Thomas Goossens

- 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 lens prescription is unknown
- This approach will be valuable for simulating **metalenses**

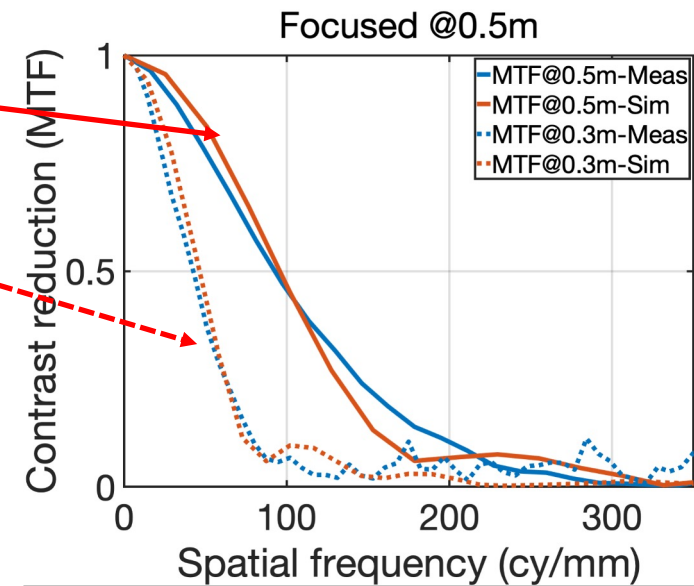
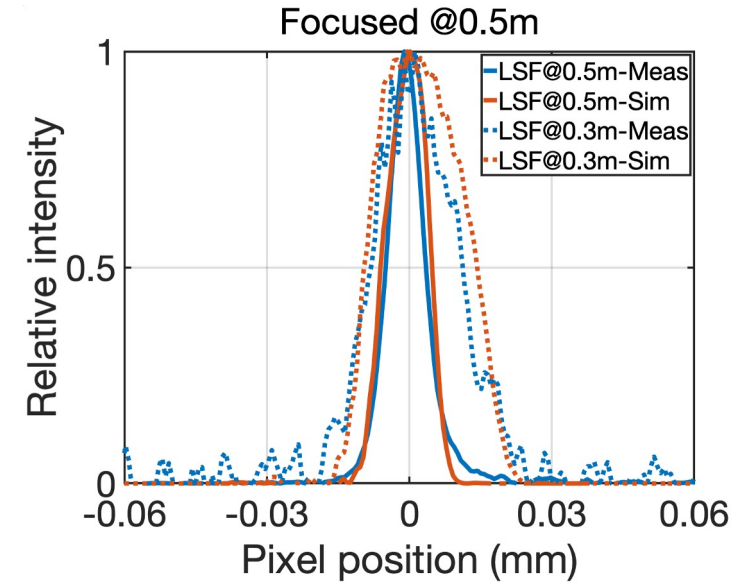
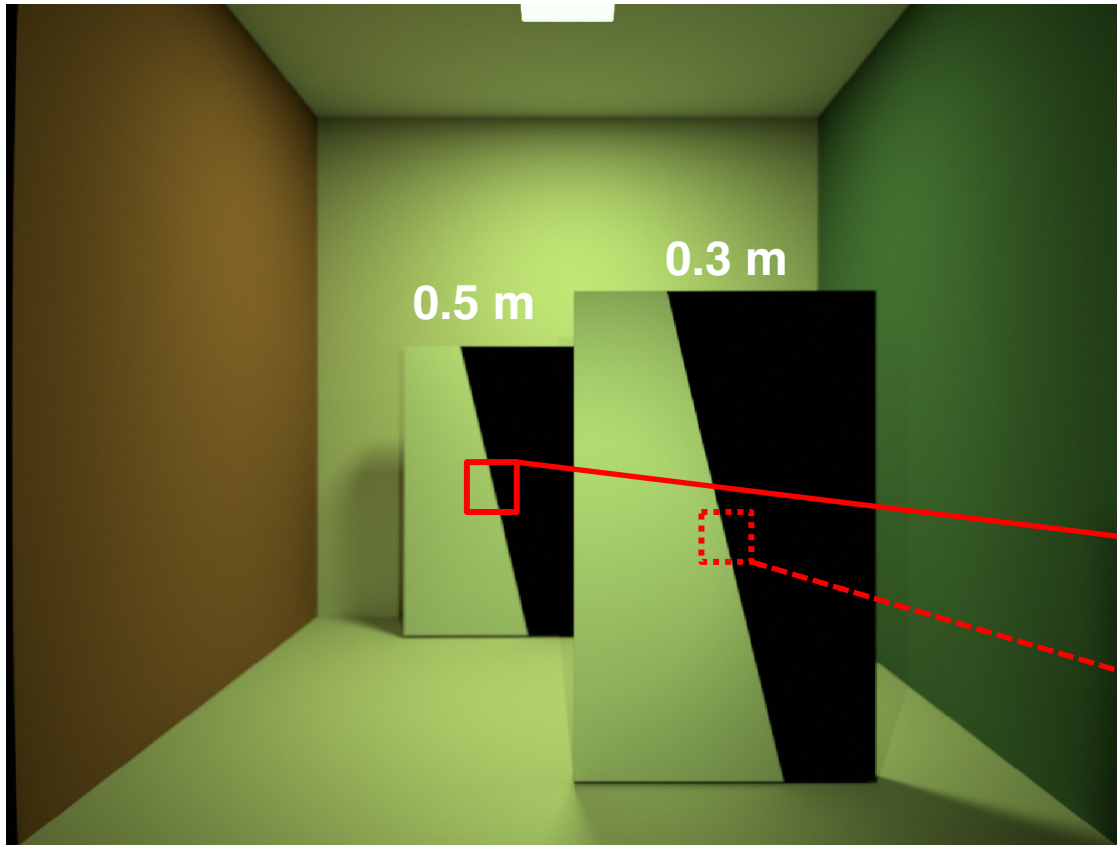


Ray-transfer function for modeling optics in camera simulation with 3D scenes (in press, Optics Express).

Goossens, Lyu, Ko, Wan, Farrell and Wandell

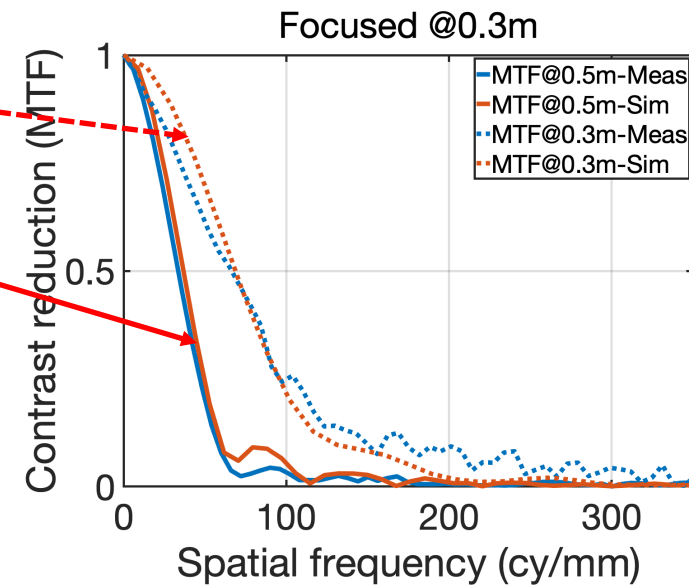
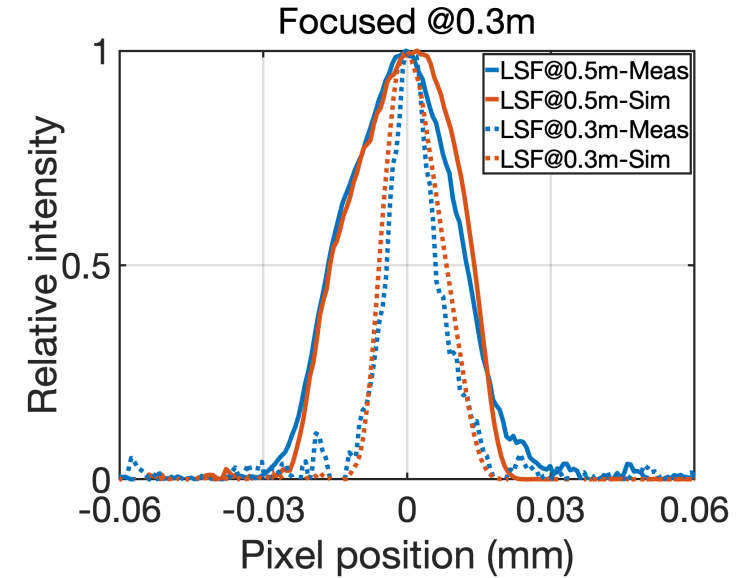
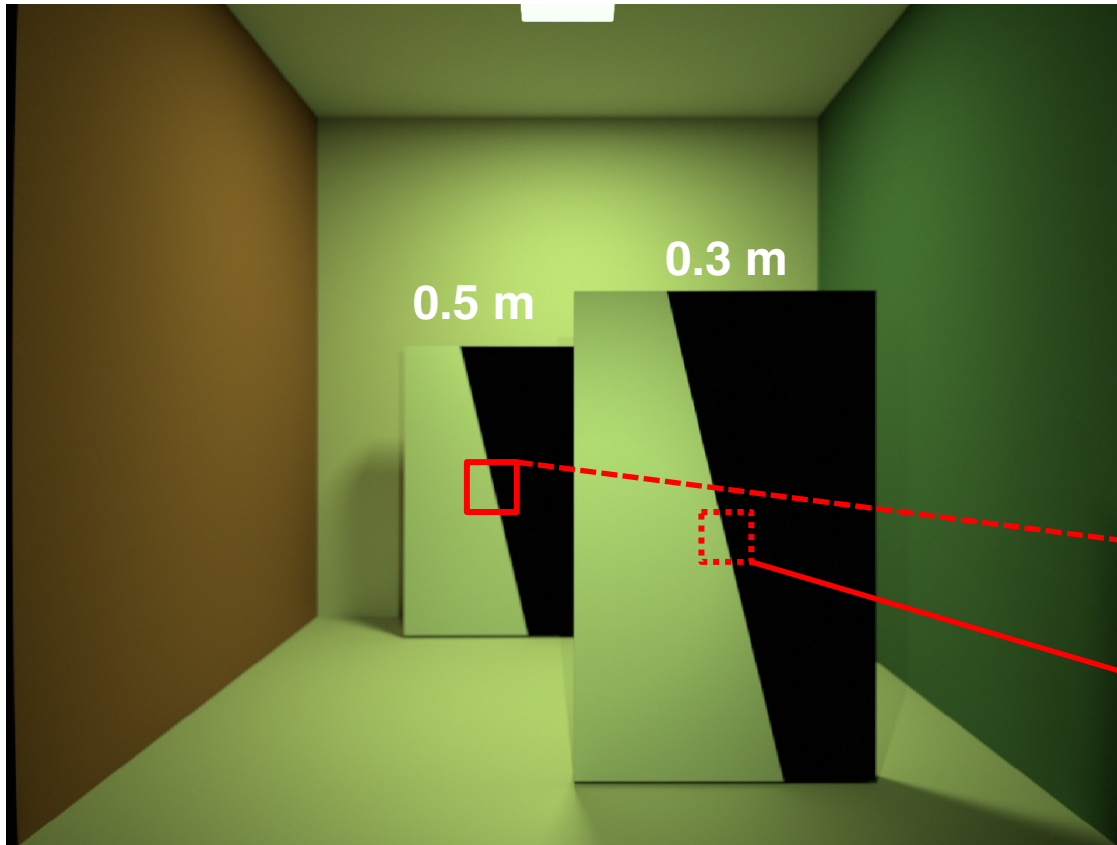
Focus-dependent LSF/MTF (depth of field)

Focused at 0.5 meters



Focus-dependent LSF/MTF (depth of field)

Focused at 0.3 meters



Vendors do not have to disclose pixel and lens design



Users can simulate more complex scenes

This approach also may value for incorporating information into the rendering about metasurfaces

Validation: sensor properties

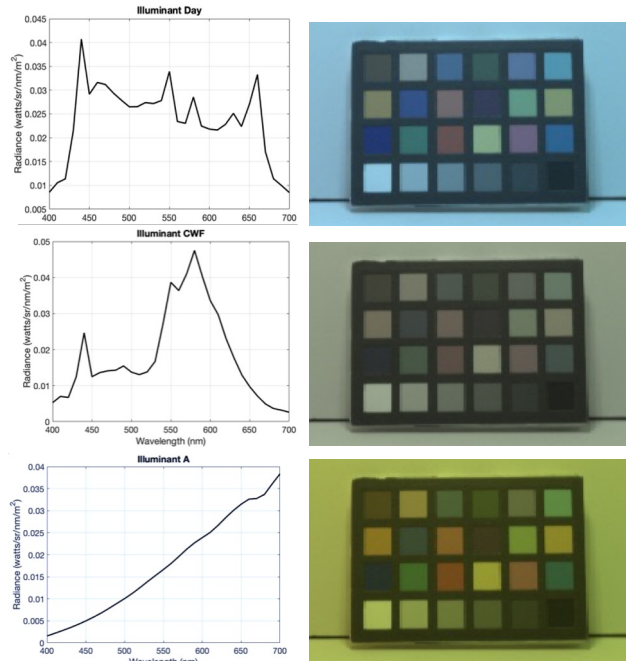
- **Sensor:** The Google Pixel 4a uses the Sony IMX363 CMOS sensor
- **Electronics:** We used vendor data and also measured electronic noise properties, including read noise, dark noise, DSNU and PRNU

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)

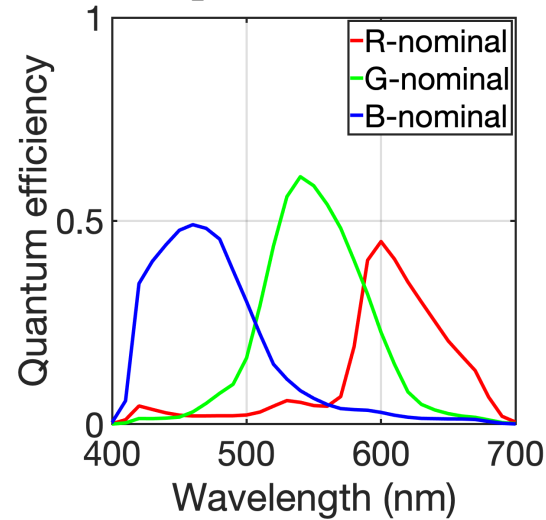


Color channel quantum efficiencies

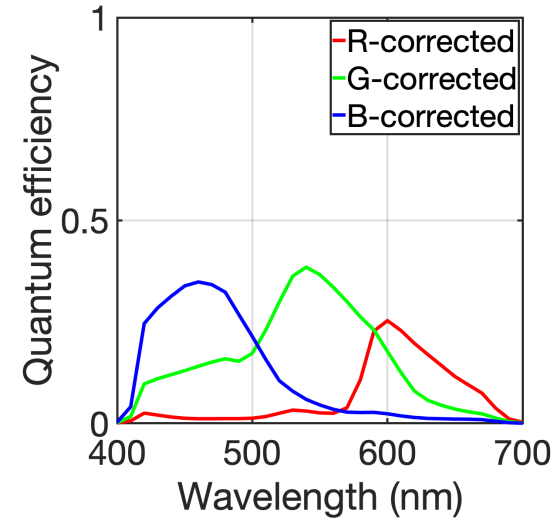
24 color patches
3 illuminants



Vendor specification

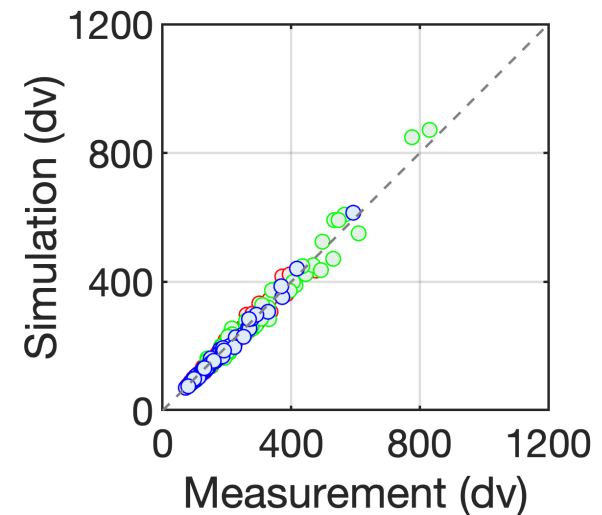
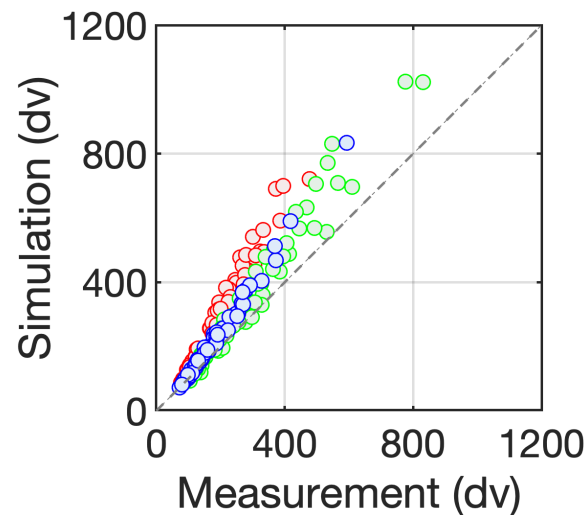


Calibration



Calibrated correction:
color channel gain and
crosstalk

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



Sensor noise measurements

A combination of electronic noise and photon noise

Many small uniform regions

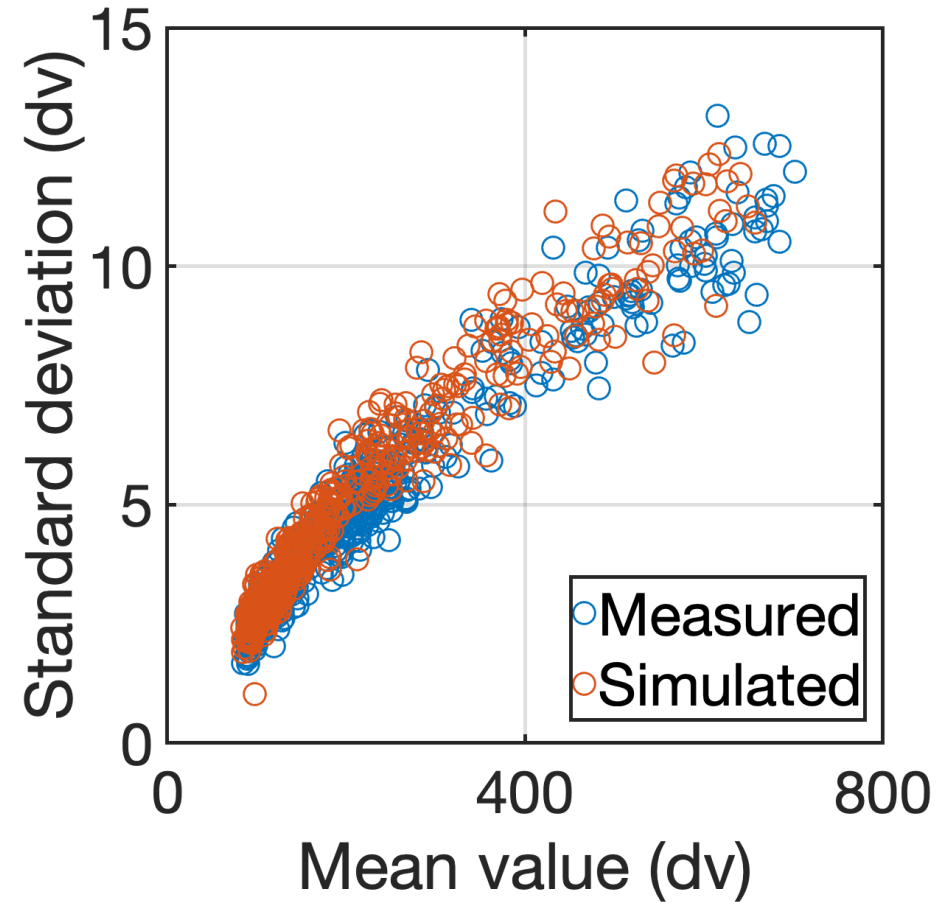
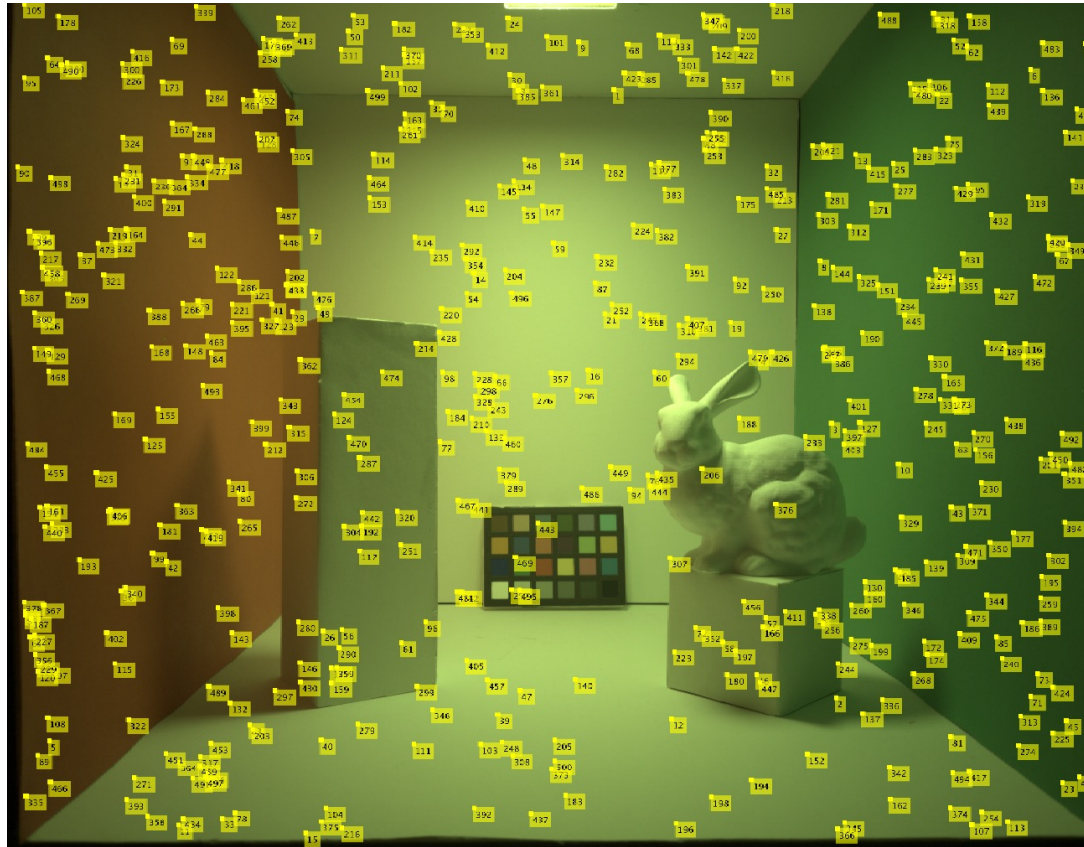
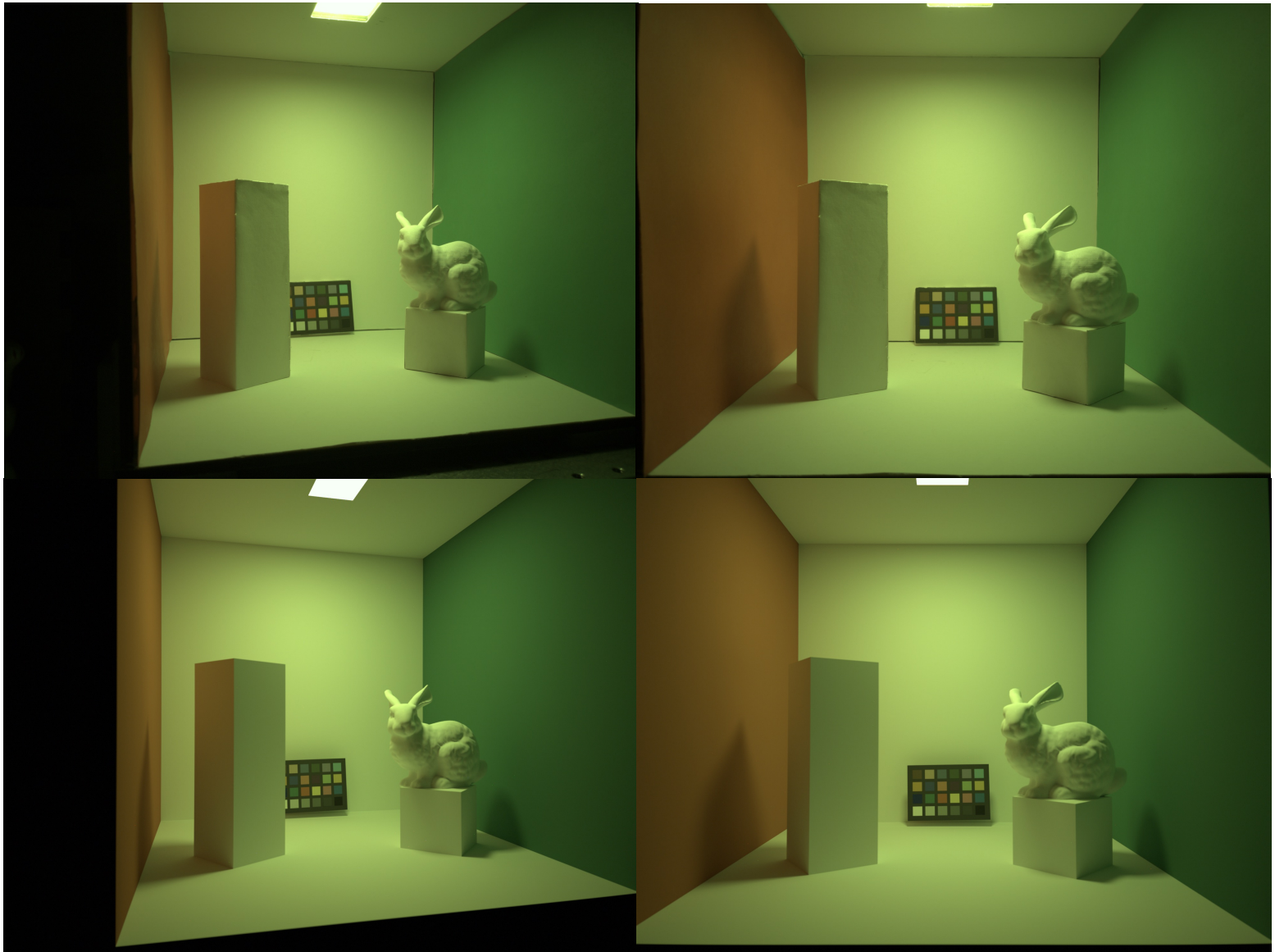


Image system simulation evaluation using 3D scenes

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





Zhenyi Liu

A system for generating complex physically accurate sensor images for automotive applications

Zhenyi Liu^{1,2}, Minghao Shen², Jiaqi Zhang³, Shuangting Liu³, Henryk Blasinski¹, Trisha Lian¹, Brian Wandell¹
1. Stanford University, 2. Jilin University, 3. Beihang University

Abstract

We describe an open-source simulator that creates sensor irradiance and sensor images of typical automotive scenes in urban settings. The purpose of the system is to support camera design and testing for automotive applications. The user can specify scene parameters (e.g., scene type, road type, traffic density, time of day)

distributions that enable us to model the impact of wavelength-dependent components, including the optics and sensors (Blasinski et al. 2018).

This paper describes an open-source and freely distributed toolbox to synthesize scene spectral radiances and sensor data for neural network automotive applications. The software includes procedural methods to generate a large number and variety of scenes from graphics assets stored in a database. The software simulates

Neural Network Generalization: The Impact of Camera Parameters

ZHENYI LIU^{1,2}, TRISHA LIAN², JOYCE FARRELL², AND BRIAN A. WANDELL²

¹State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 13000, China

²Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA

Corresponding author: Zhenyi Liu (zhenyiliu27@gmail.com)

This work was supported by the Jilin University.

ABSTRACT We quantify the generalization of a convolutional neural network (CNN) trained to identify cars. First, we perform a series of experiments to train the network using one image dataset - either synthetic or from a camera - and then test on a different image dataset. We show that generalization between

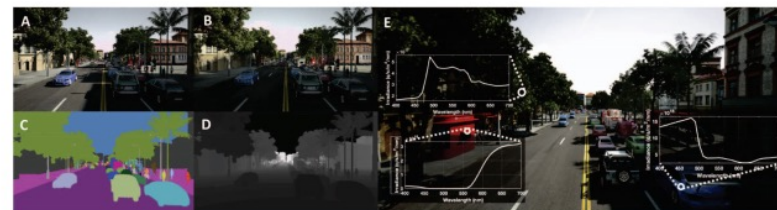


This ICCV Workshop paper is the Open Access version, provided by the Computer Vision Foundation. Except for this watermark, it is identical to the accepted version; the final published version of the proceedings is available on IEEE Xplore.

Soft Prototyping Camera Designs for Car Detection Based on a Convolutional Neural Network

Zhenyi Liu^{1,2}, Trisha Lian¹, Joyce Farrell¹, and Brian Wandell¹

¹Stanford University, USA, ²Jilin University, China
{zhenyiliu, tlian, jefarrel, wandell}@stanford.edu



ISETAuto: Detecting vehicles with depth and radiance information

ZHENYI LIU¹, JOYCE FARRELL² AND BRIAN WANDELL²

¹State Key Laboratory of Automotive Simulation and Control, Jilin University (e-mail: zhenyiliu27@gmail.com)

²Stanford University (e-mail: jefarrel, wandell@stanford.edu)

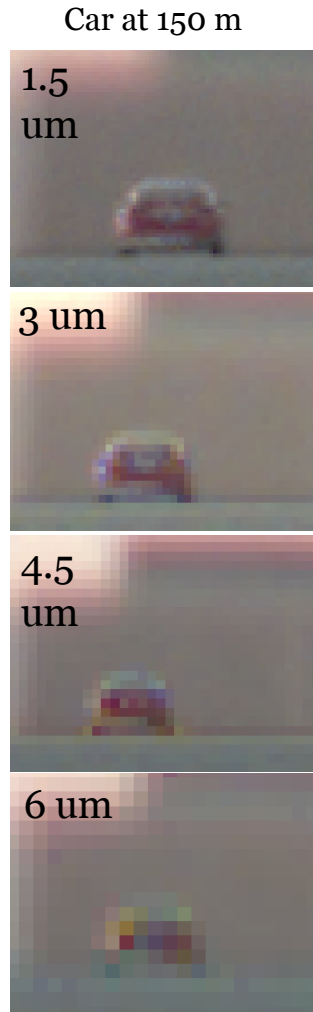
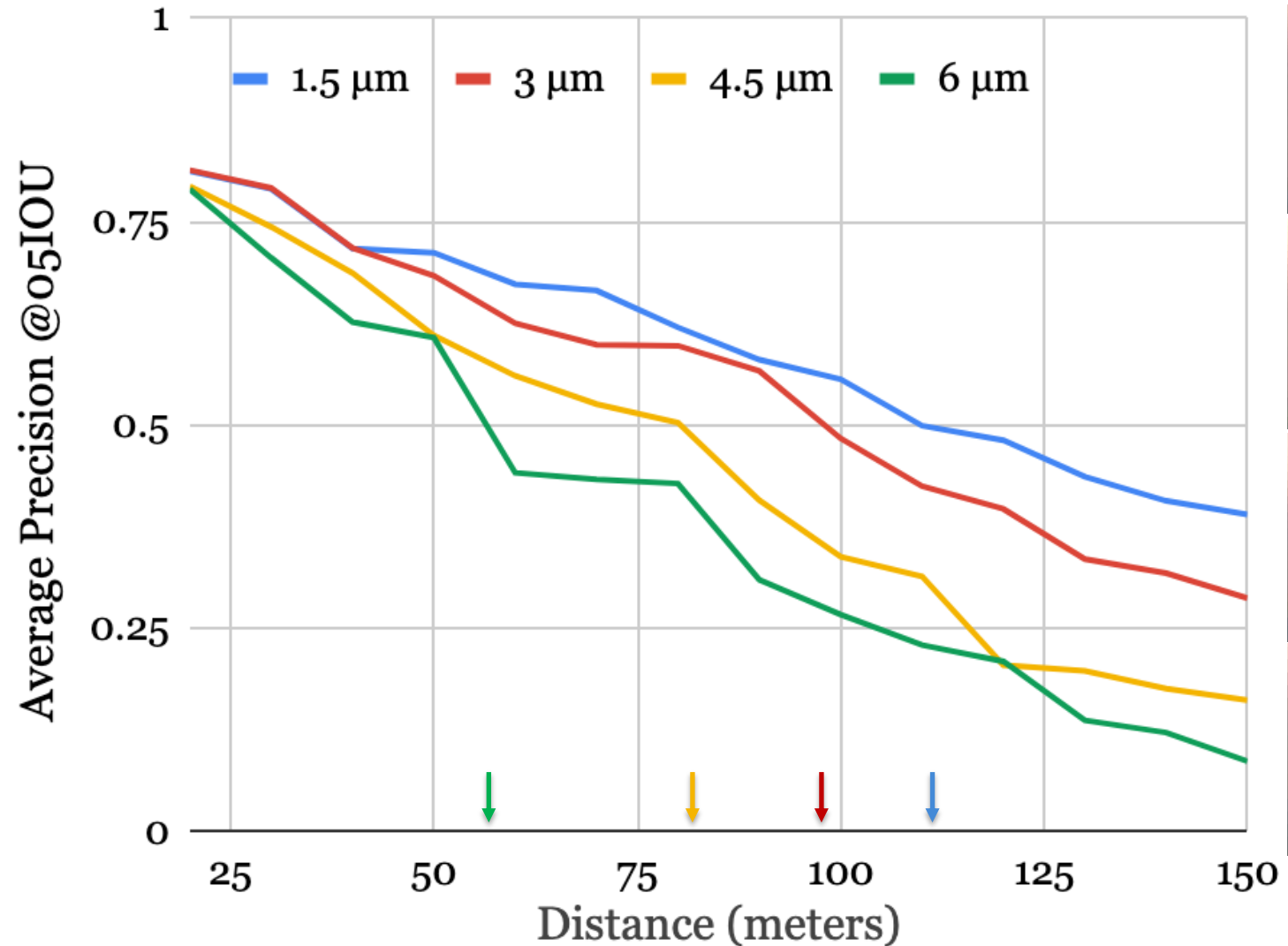
Corresponding author: Zhenyi Liu (e-mail: zhenyiliu27@gmail.com)

Supported by Jilin University. We thank Boyd Fowler at Omnivision and Sergio Goma at Qualcomm for drawing our attention to prior work on RGB-D sensor technology.

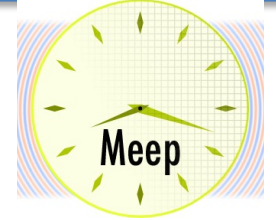
ABSTRACT Autonomous driving applications use two types of sensor systems to detect vehicles - depth sensing LiDAR and radiance sensing cameras. We compare the performance (average precision) of a ResNet for vehicle detection in complex, daytime, driving scenes when the input is a depth map $[D = d(x,y)]$, a

Meaningful performance metrics computed from simulation

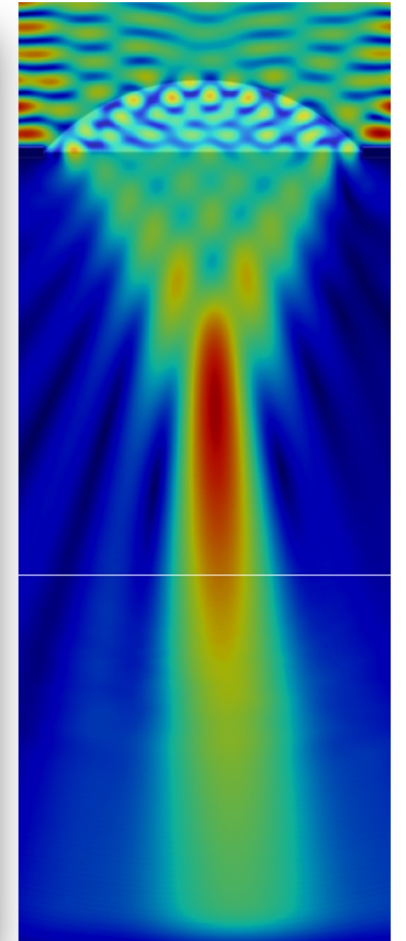
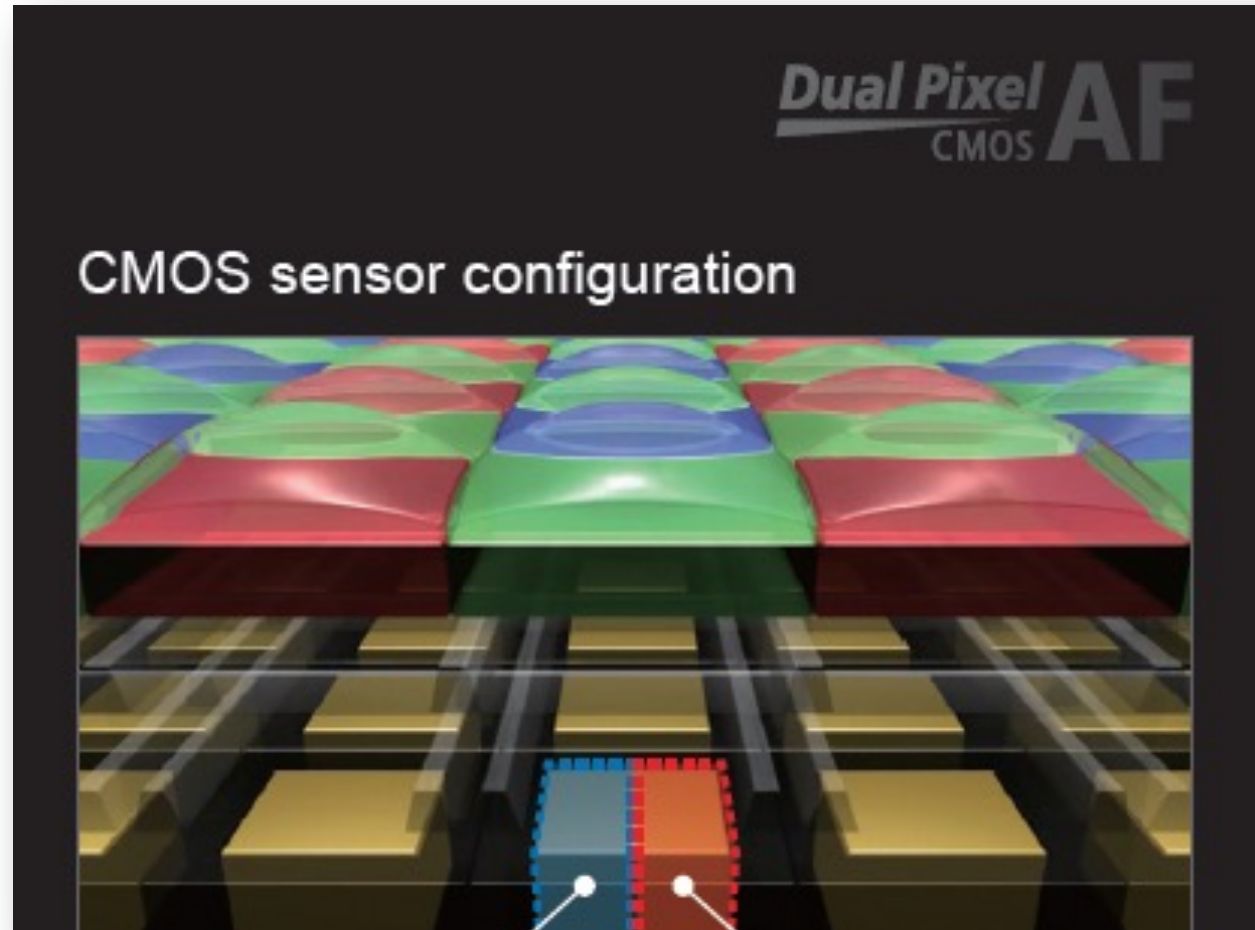
- This analysis simulates average precision of a CNN (ResNet-50) comparing with cameras with different pixel sizes
- The accuracy depends on pixel size, systematically
- The simulation includes meaningful units
- The performance of the 1.5 μm pixels is similar to current LIDAR detectors



What is next? Wave properties, light field designs, camera controllers



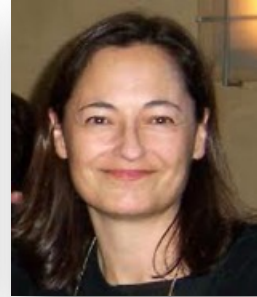
- Dual pixel (Canon) for autofocus
- Quad pixel and more for single-shot depth
- **Technology has scaled to submicron pixels:** Wave optics matters, and a simulation technology that integrates with ray optics is necessary



What goes into an image systems simulation for the human eye



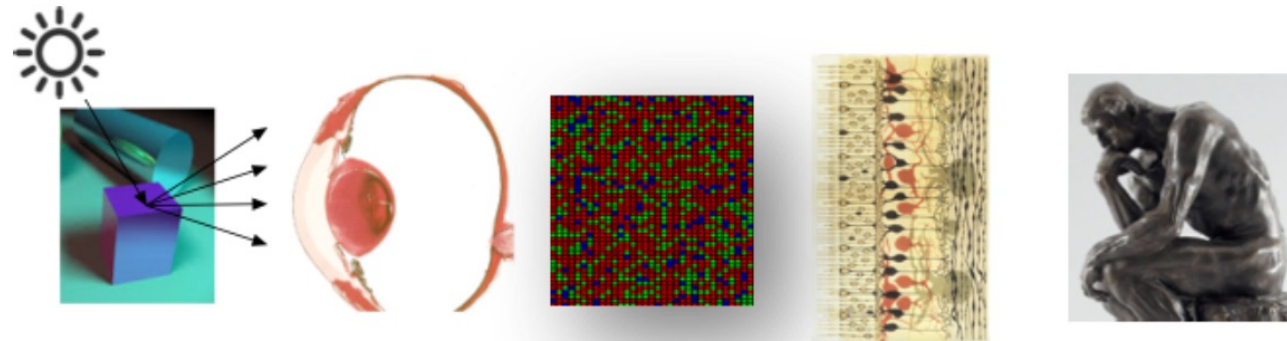
Nicolas Cottaris



Joyce Farrell



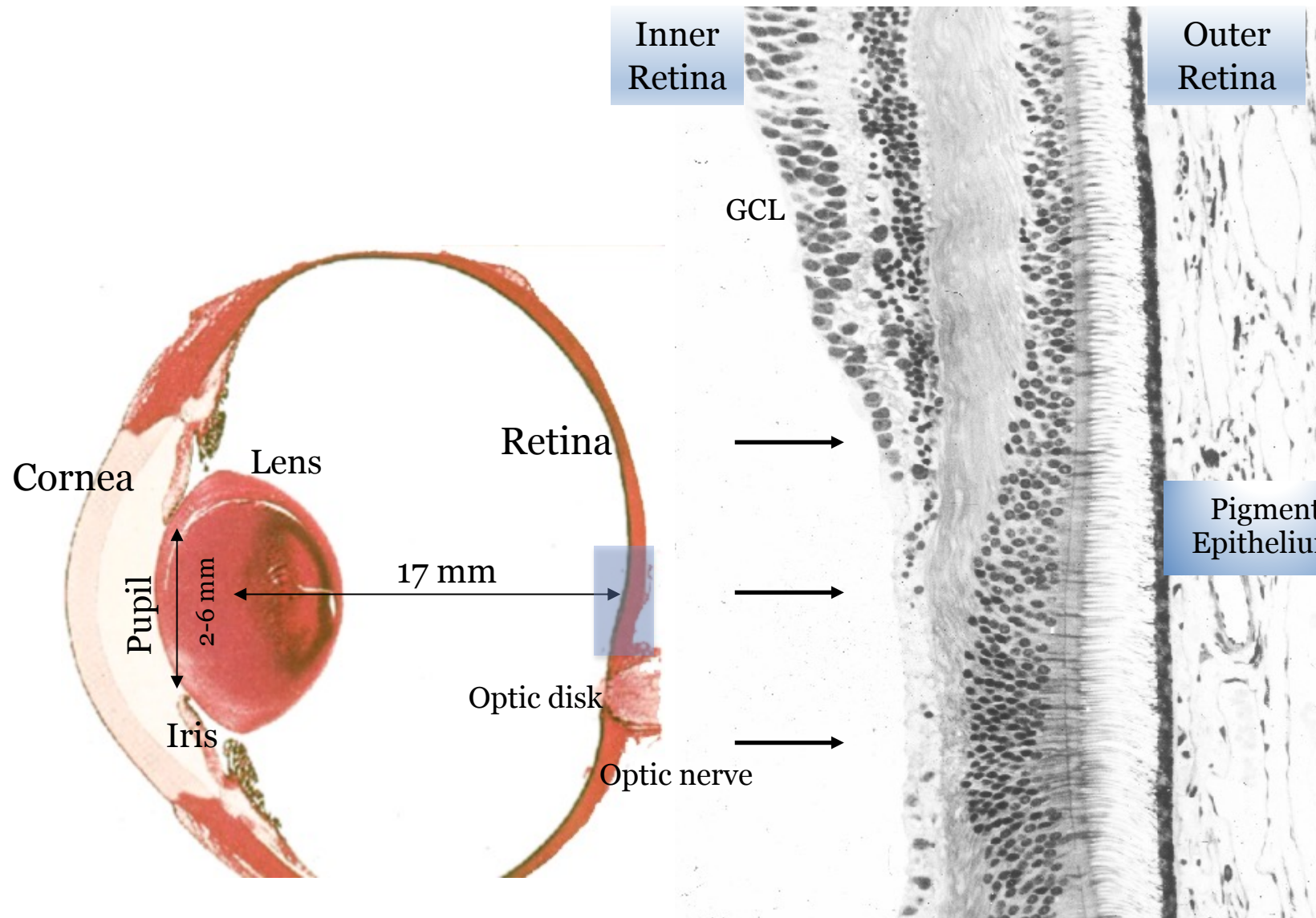
David Brainard



<https://github.com/isetbio/isetbio/wiki>

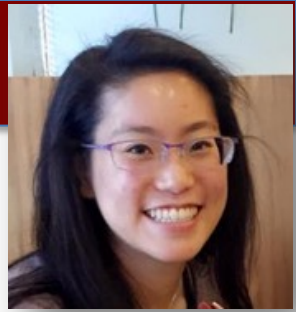
The Image System Engineering Toolbox for Biology (ISETBio) supports vision science calculations. The Matlab software enables you to create spectral radiance scenes and use these scenes as input to estimate the effects of human optics, eye movements, cone absorptions and photocurrent, and retinal cell properties. We hope to expand the software to model visual responses in thalamus and cortex. We have recorded some that introduce various aspects of ISETBio, and also that introduce basic vision science concepts. These may be found under the Videos menu item to the right.

The primate fovea (pit) contains mainly cones and is specialized for high acuity and color



- 5 x 5 cm, 0.4 mm thick
- 5×10^6 cones
- 10^8 rods
- Foveal cone: $1 \mu\text{m}^2$
- Contacts per cone: 250
- 10^6 optic nerve fibers

Comparison of eye models: 3D renderings

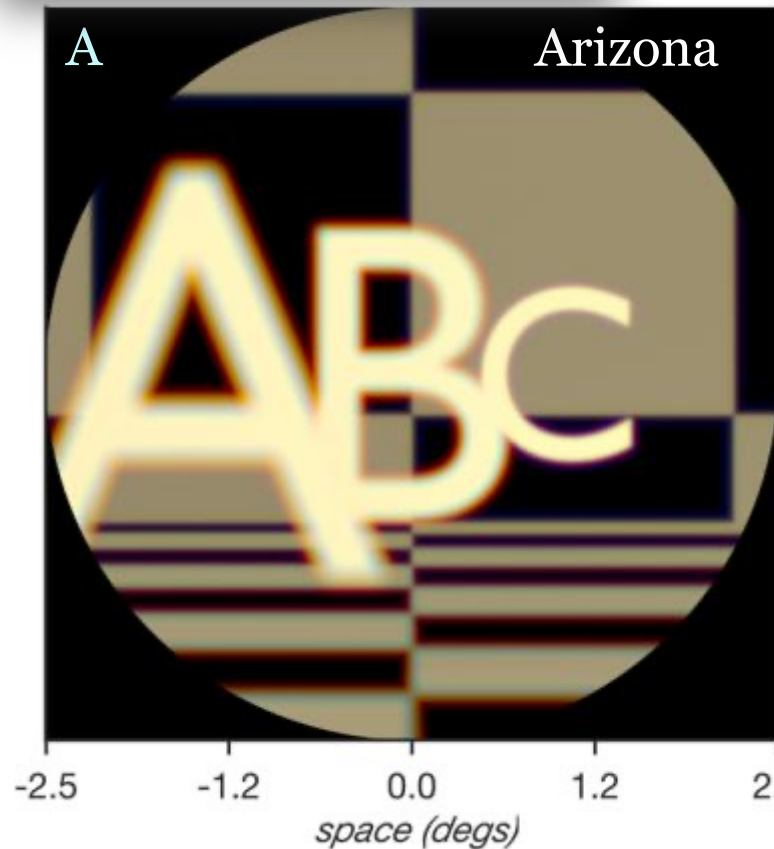


Trisha Lian

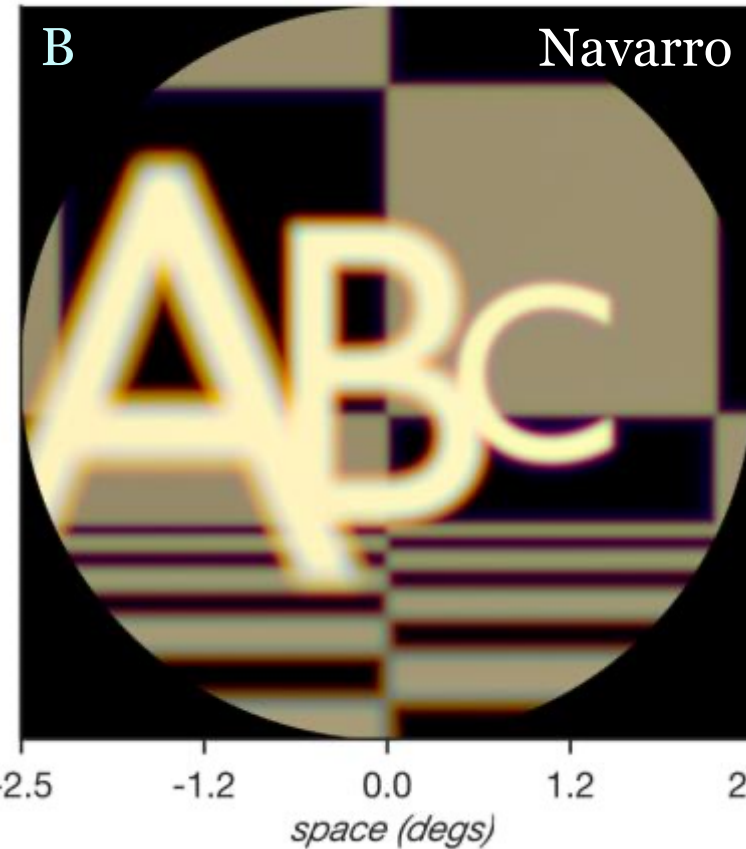
The PBRT-ISET3d code implements three major human eye models
(Lian et al. 2019, Journal of Vision).

Remember: these images represent underlying spectral irradiance

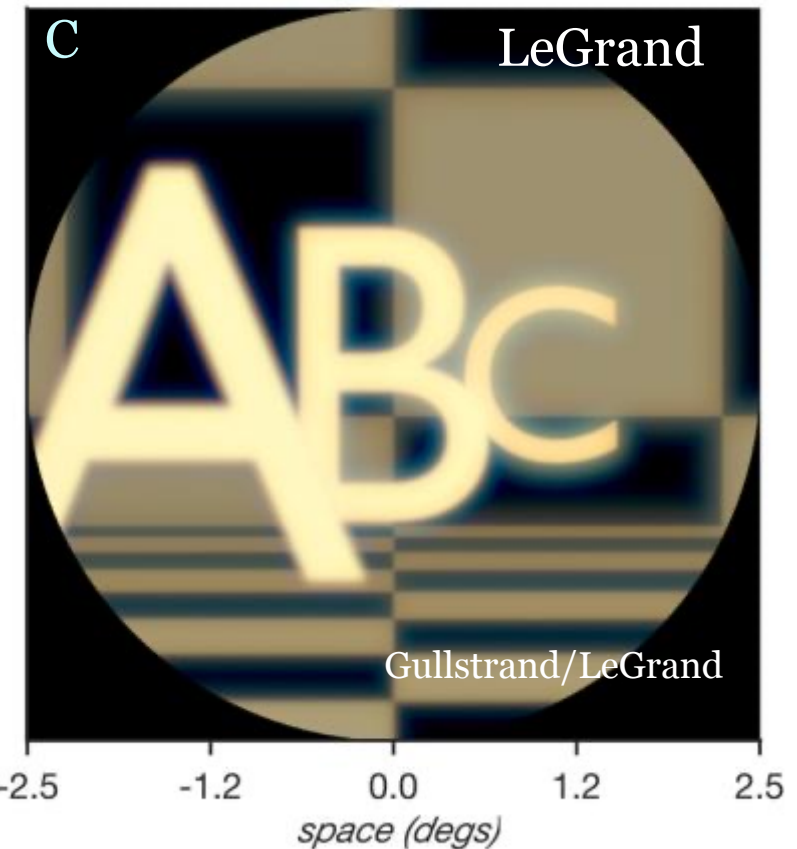
OPEN ACCESS
Article | October 2019
Ray tracing 3D spectral scenes through human optics models
Trisha Lian; Kevin J. MacKenzie; David H. Brainard; Nicolas P. Cottaris; Brian A. Wandell
[+ Author Affiliations](#)
Journal of Vision October 2019, Vol.19, 23. doi:<https://doi.org/10.1167/19.12.23>



Schwiegerling, et al. 2004



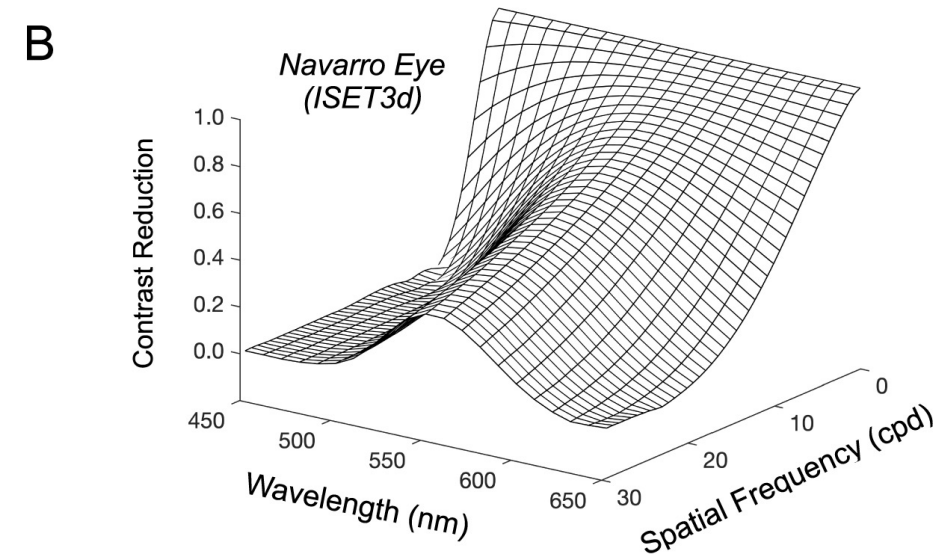
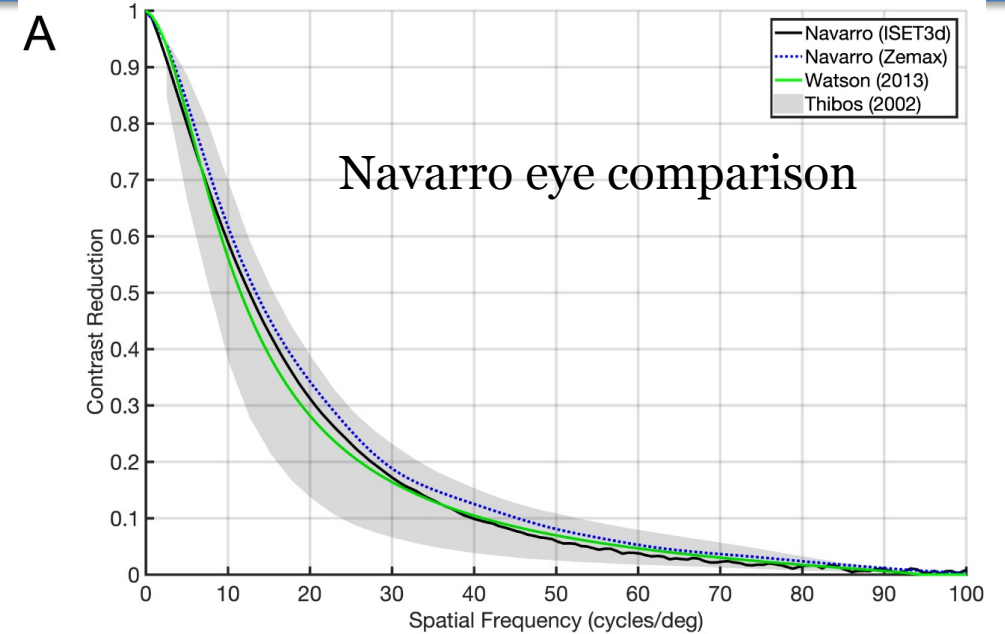
Escudero-Sanz & Navarro, 1999



Hage & Le Grand, 1980

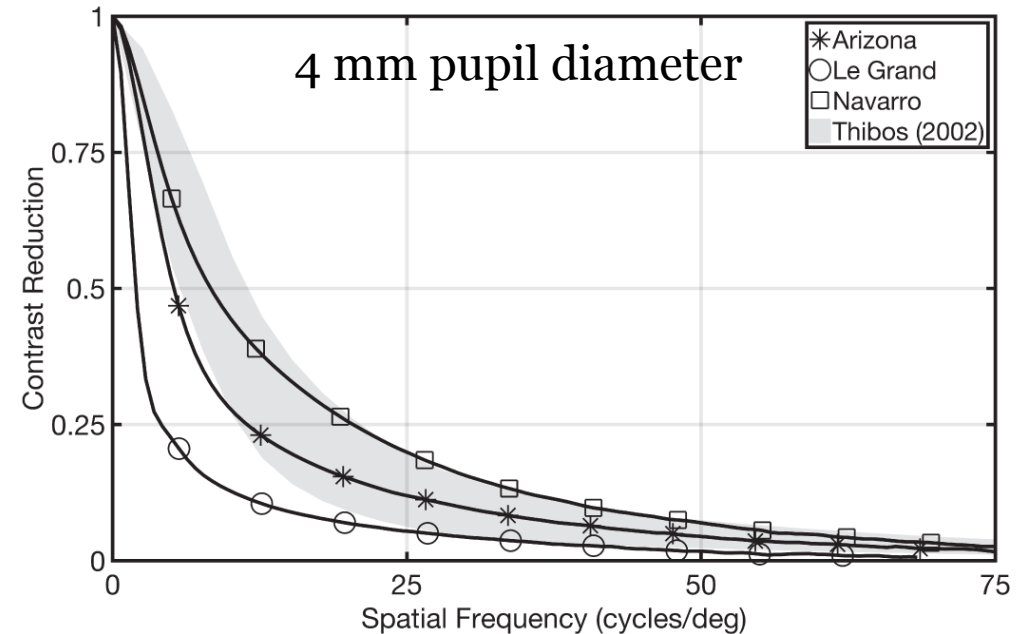
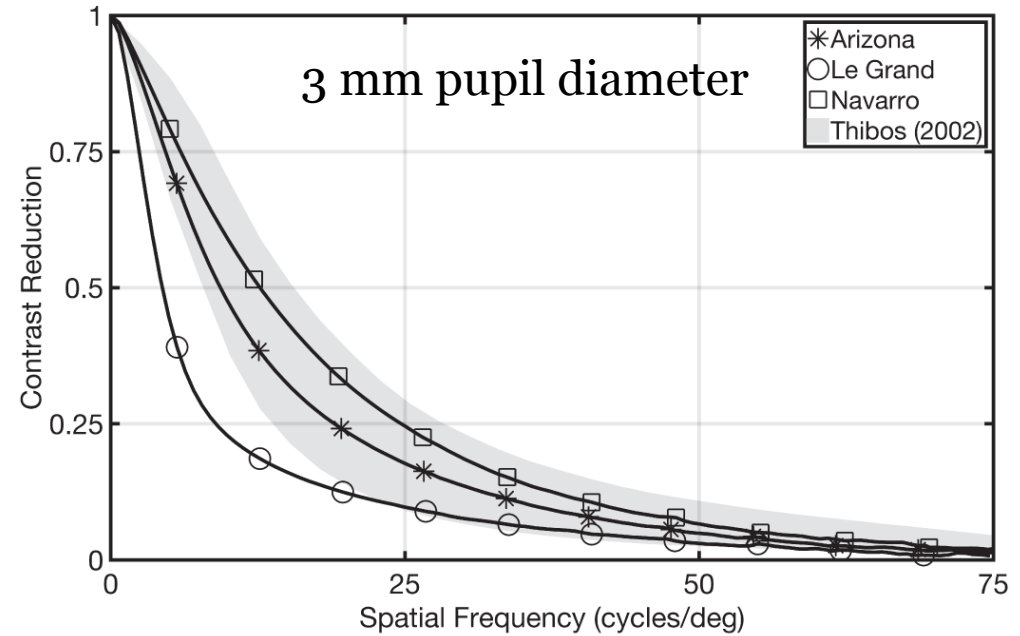
Validation of the ray tracing methods

- The PBRT implementation matches the Zemax calculations for these models
- We compared with L. Thibos' wavefront aberration data from many eyes (shaded) – which is the basis of the Watson (green) CSF summary as well
- We calculated as a function of wavelength, too – note huge chromatic aberration

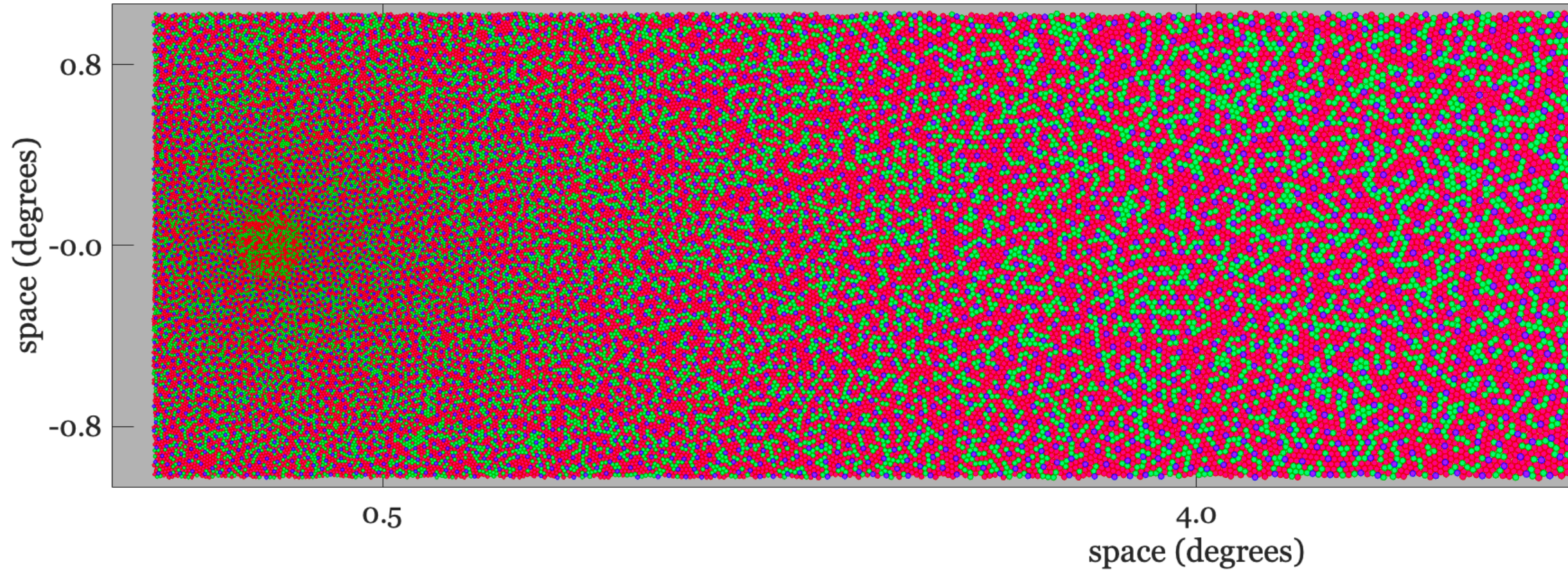


Human eye model comparisons

- There are quantitative differences between the models
- The LeGrand extension of the Gullstrand eye – is out of compliance with modern MTF measurements
- We can implement a range of eye parameters to match population distributions

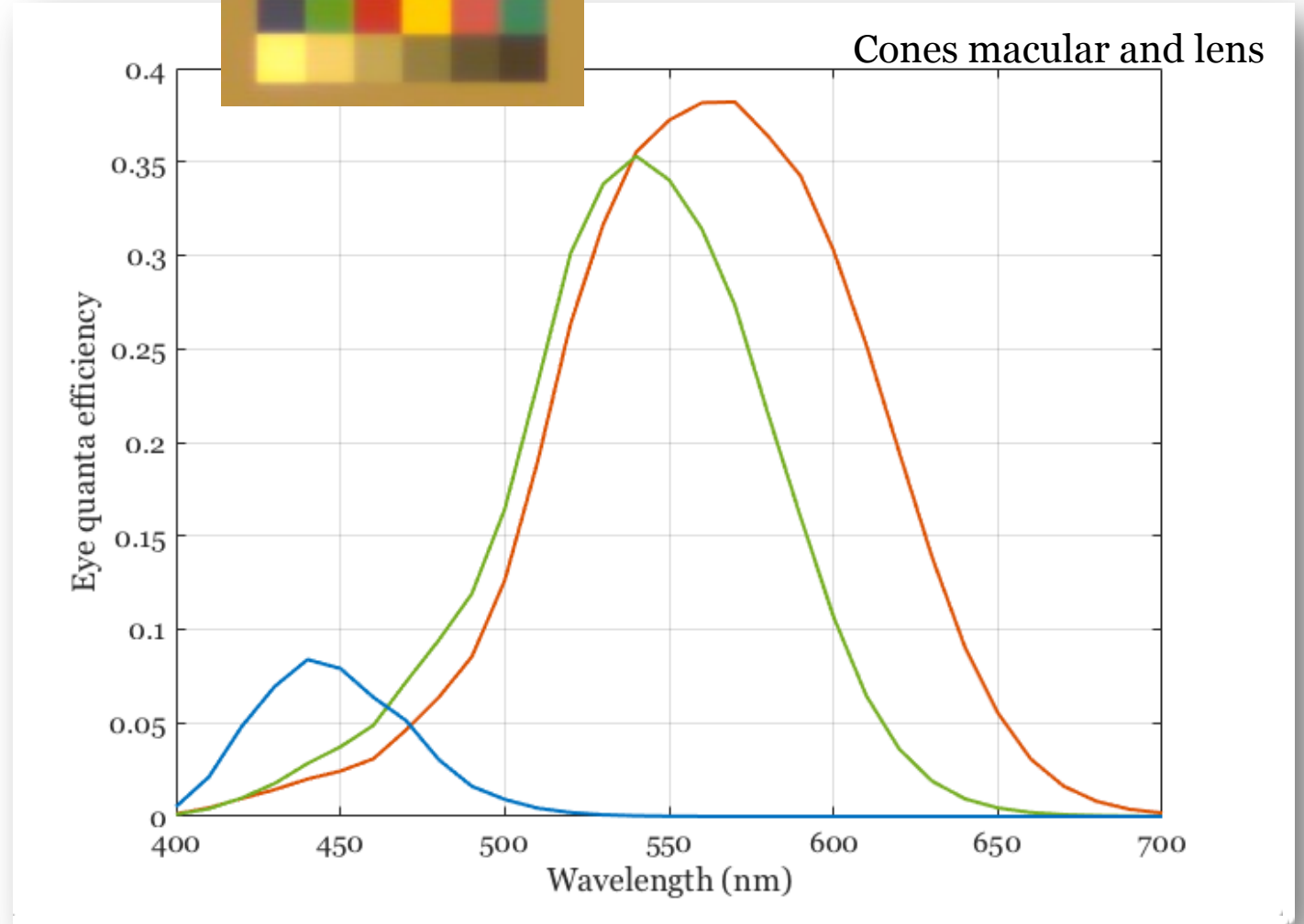
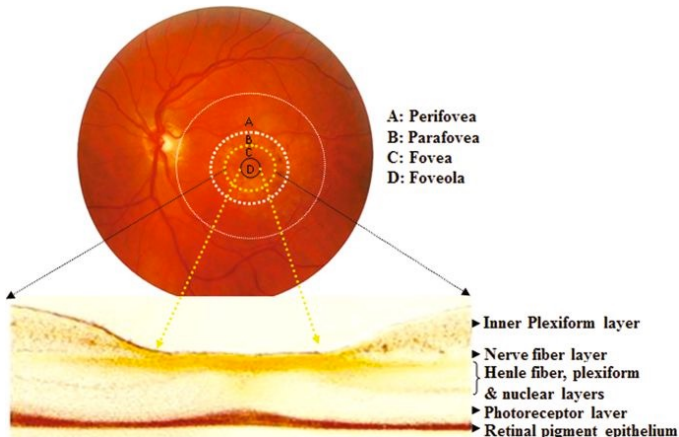


The human retina is very inhomogeneous: Cone apertures vary more than 10x



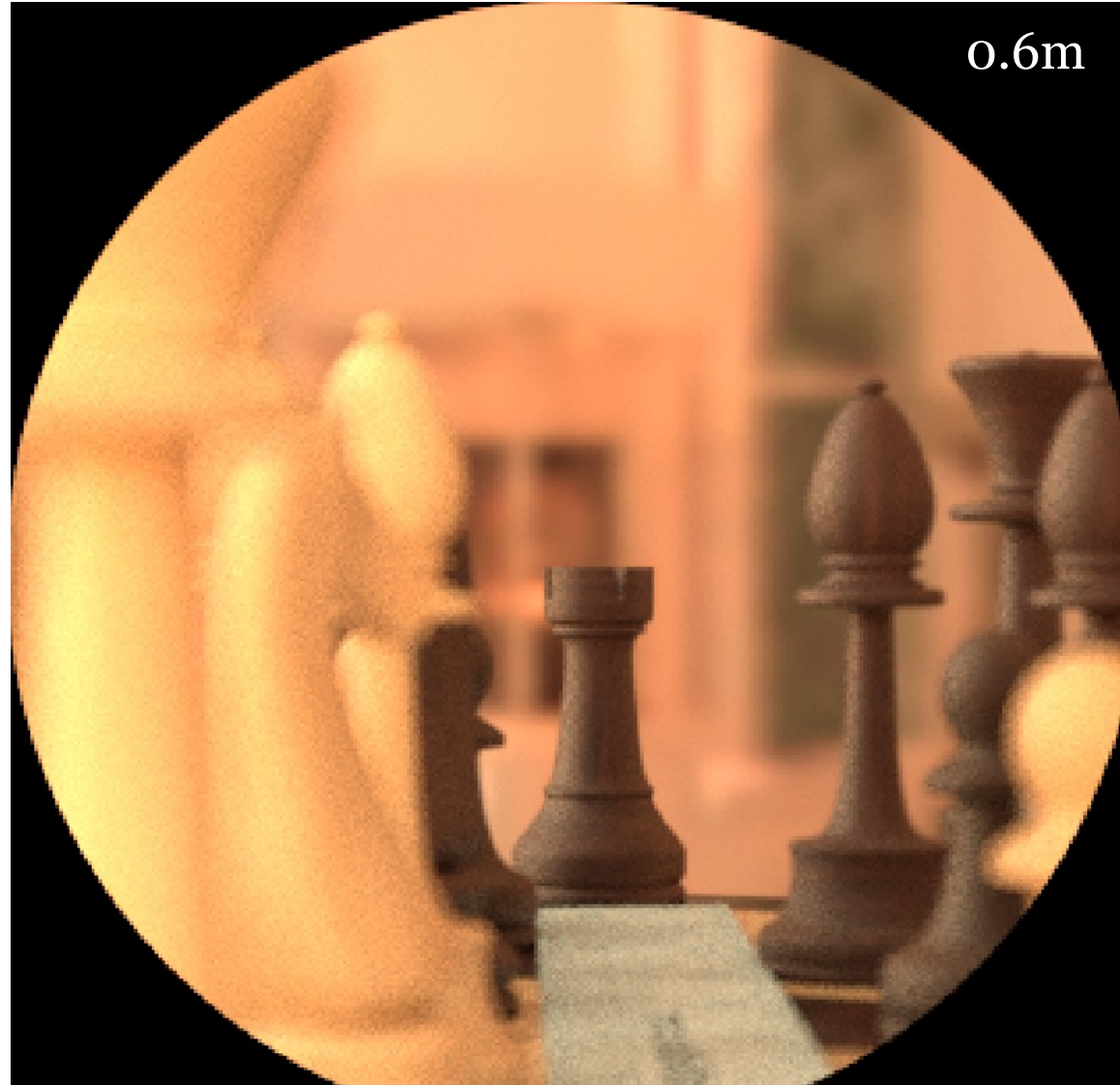
Human wavelength absorptions depend on a series of factors

- Cone photopigments (L,M,S) absorption efficiency for standard pigment density
- These curves are input referenced; they include the effect of the lens and macular pigments



Depth of field and accommodation (5 mm pupil)

- The eye models enable us to model the change of accommodation (focus) for the human eye using complex scenes – or simple scenes!!
- And to calculate the excitations at the cone mosaic
- We will get to eye movements a little later

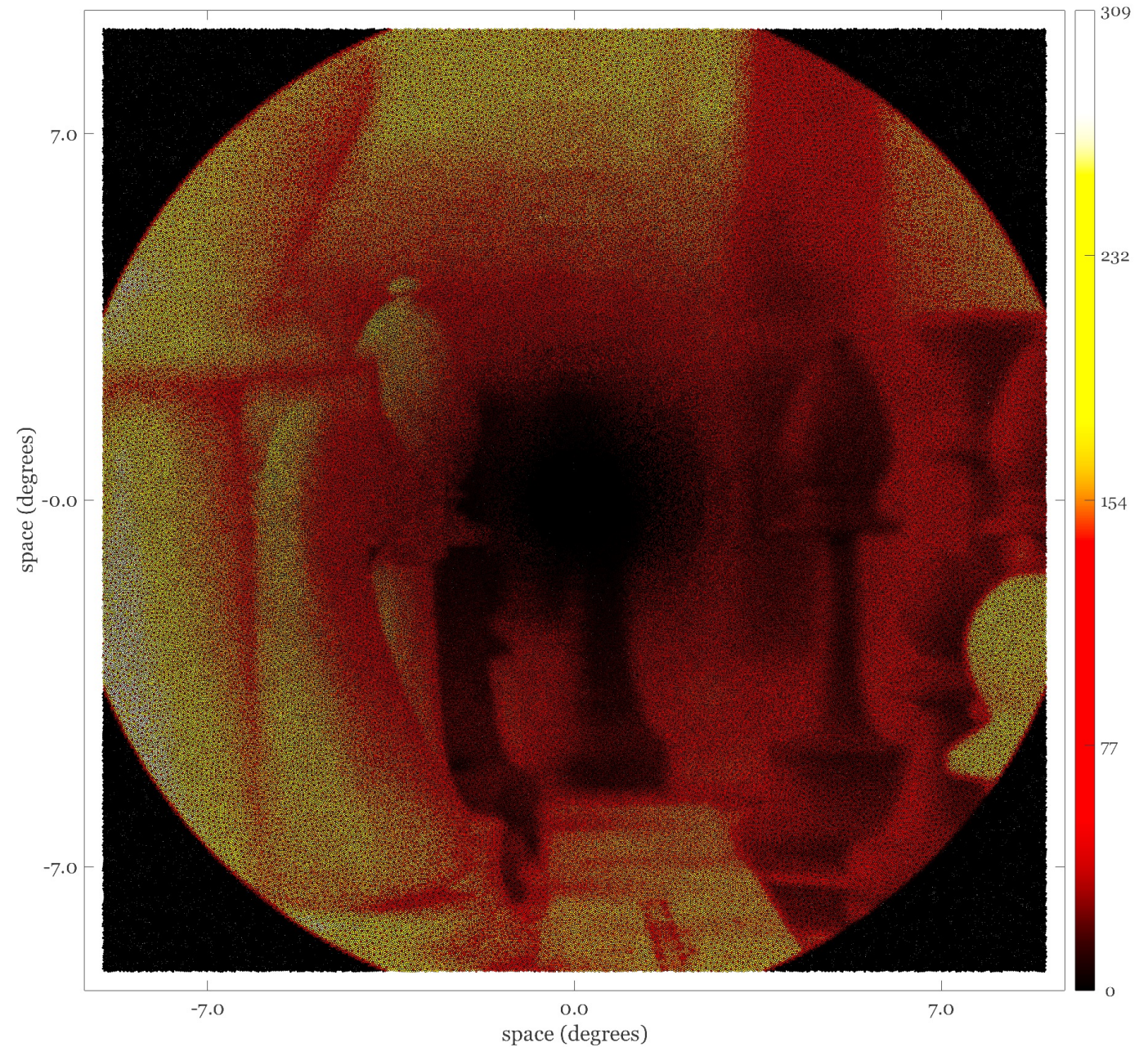


- The eye models enable us to model the change of accommodation (focus) for the human eye using complex scenes – or simple scenes!!
- And to calculate the excitations at the cone mosaic
- We will get to eye movements a little later



Cone excitations ($R^*/5$ msec)

- Notice the low number of excitations at the fovea – do you know why? (I didn't expect this before we put the system together)
- The scene mean luminance is 100 cd/m^2 which is typical of a computer monitor



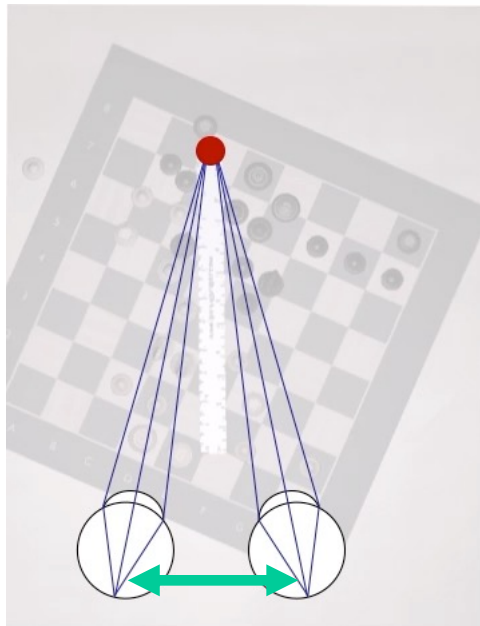
And to model vergence and accommodation

Where the eye (or eyes) is looking is controlled
thisEye.set('to',loc)

1.66 D (Left)



1.66 dpt (Right)



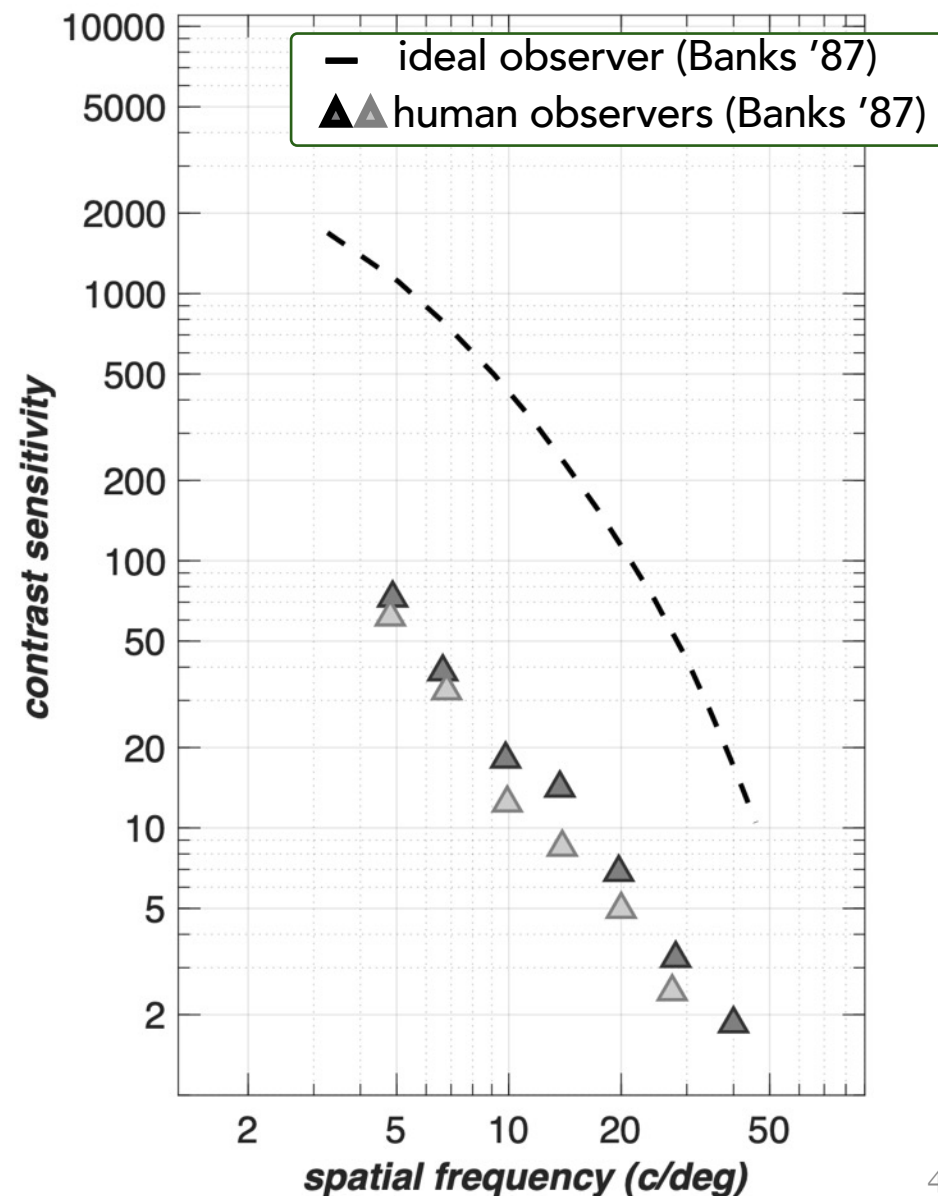
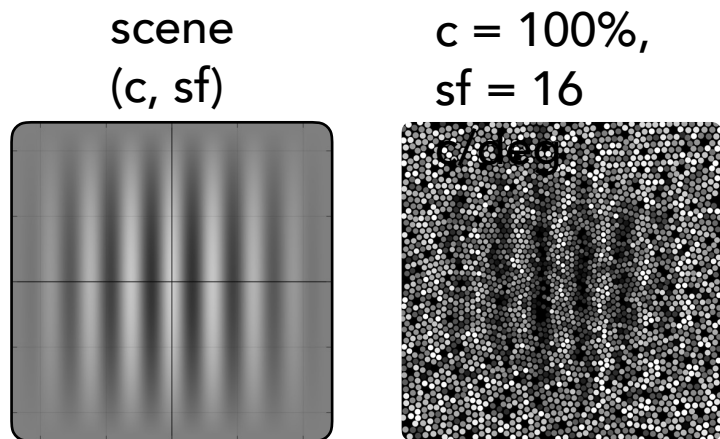
64 mm

Remember: these images represent underlying spectral irradiance

Putting it together: Analyzing the system spatial frequency contrast sensitivity

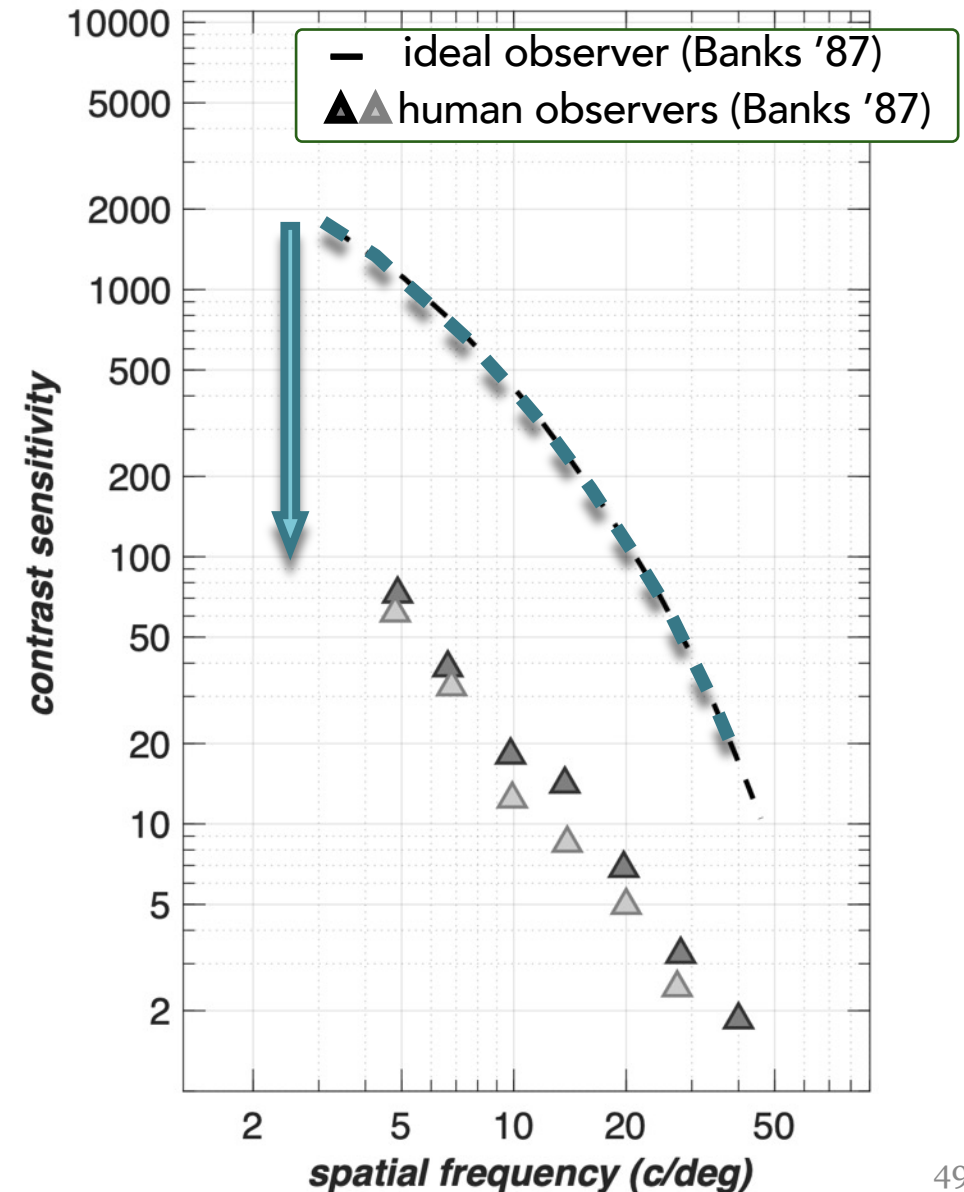
Classic work: Ideal observer accounts for shape of CSF
roll-off at high SF

- The original work from Banks et al. (1987) compared the high frequency roll-off predicted using an ideal observer and measured with a few real observers
- The predictions were based on formulae and various simplifying assumptions about the mosaic



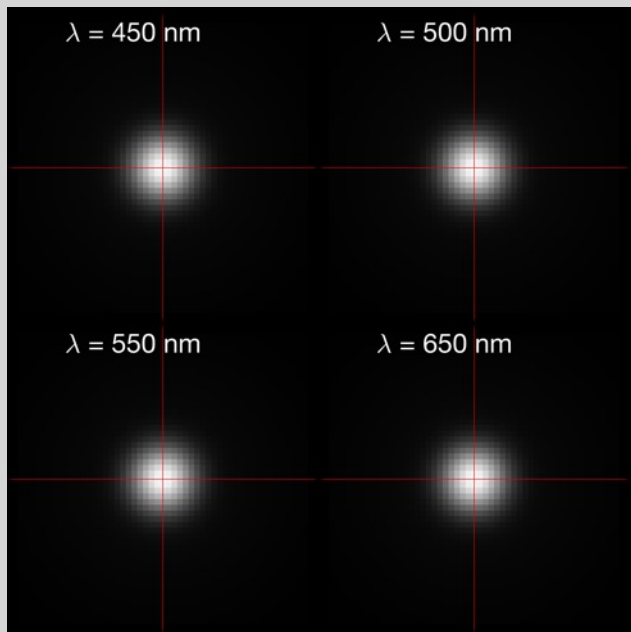
Ideal observer accounts for shape of CSF roll-off at increasing SF

- The original work from Banks et al. (1987) compared the high frequency roll-off predicted using an ideal observer and measured with a few real observers
- The predictions were based on formulae and various simplifying assumptions about the mosaic
- The shapes were in good alignment

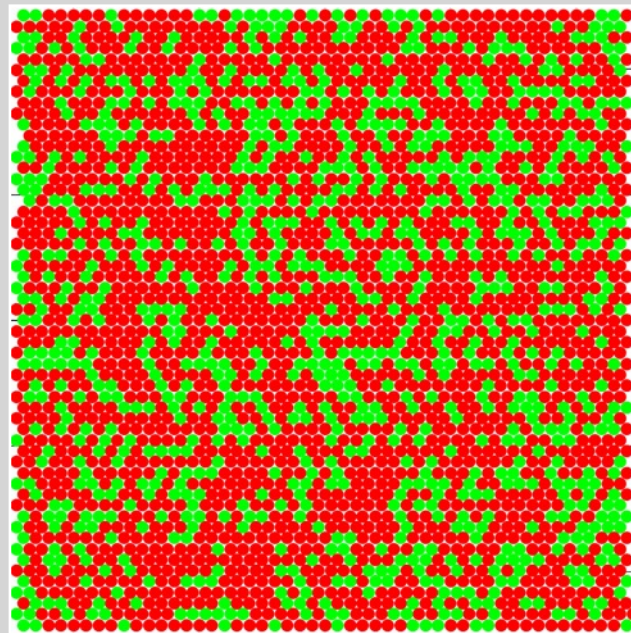


ISSETBio validation: reproduce earlier ideal observer calculation

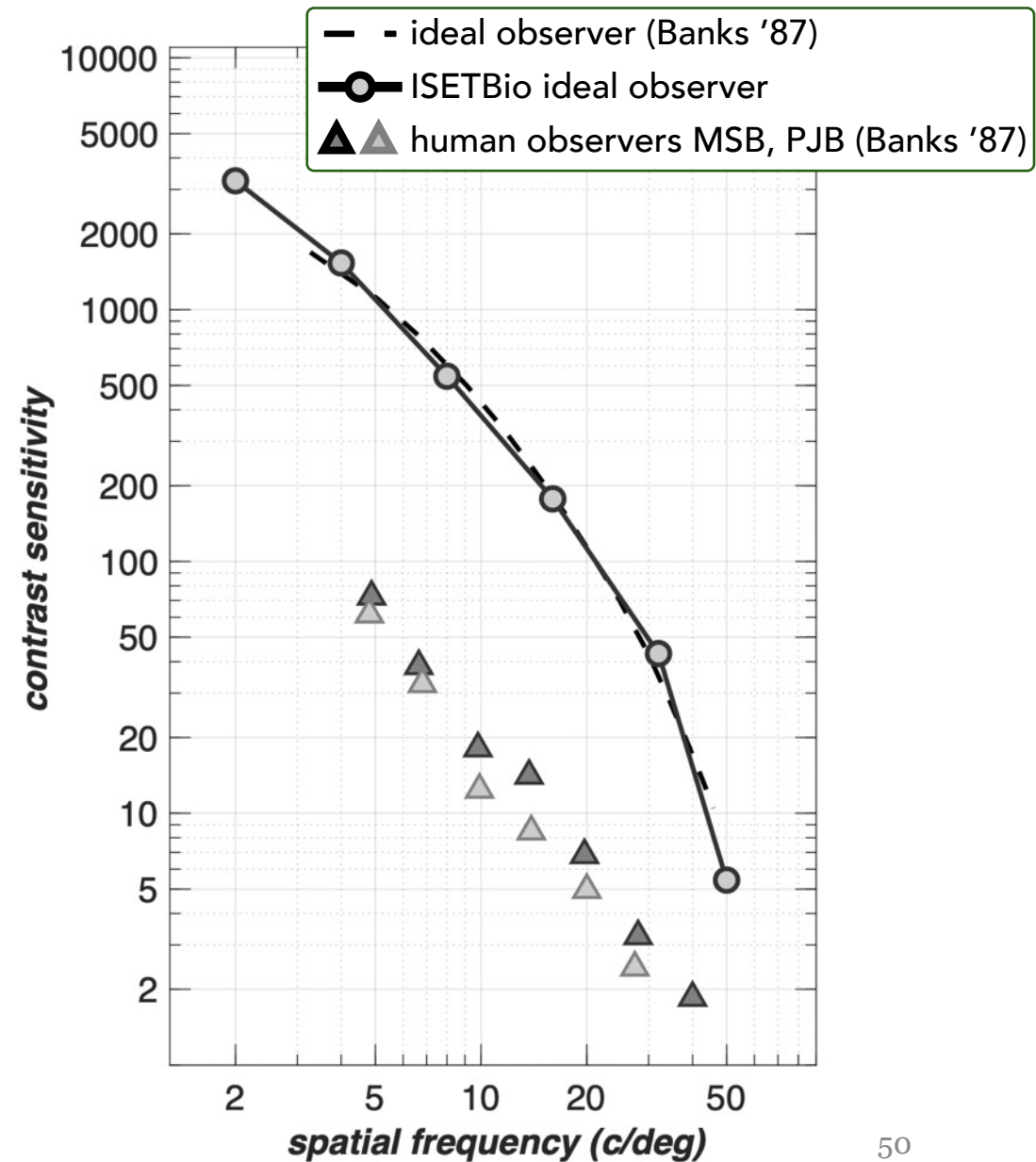
point spread function -
based optics



constant cone
density mosaics

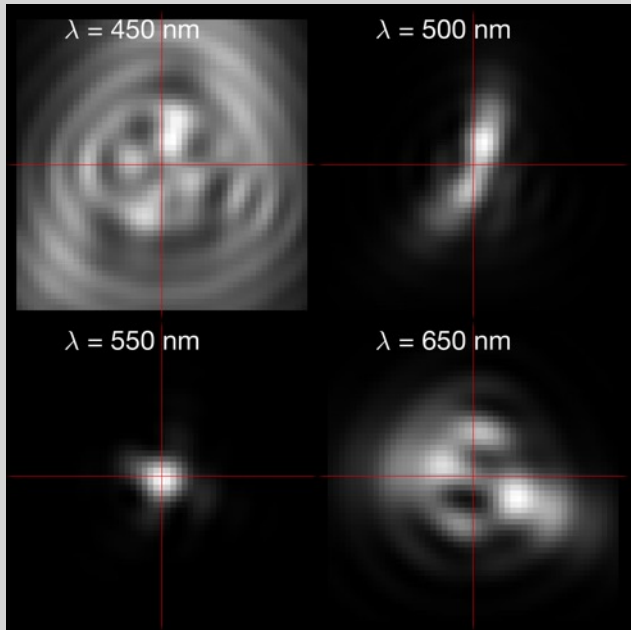


(based on Campbell & Gubisch '66)



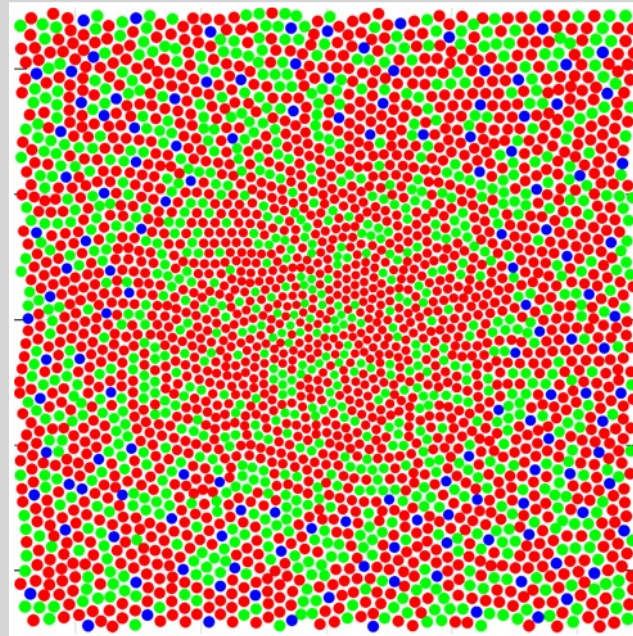
Accounting for absolute sensitivity: modern estimates of optics/mosaic

Wavefront aberration - based optics

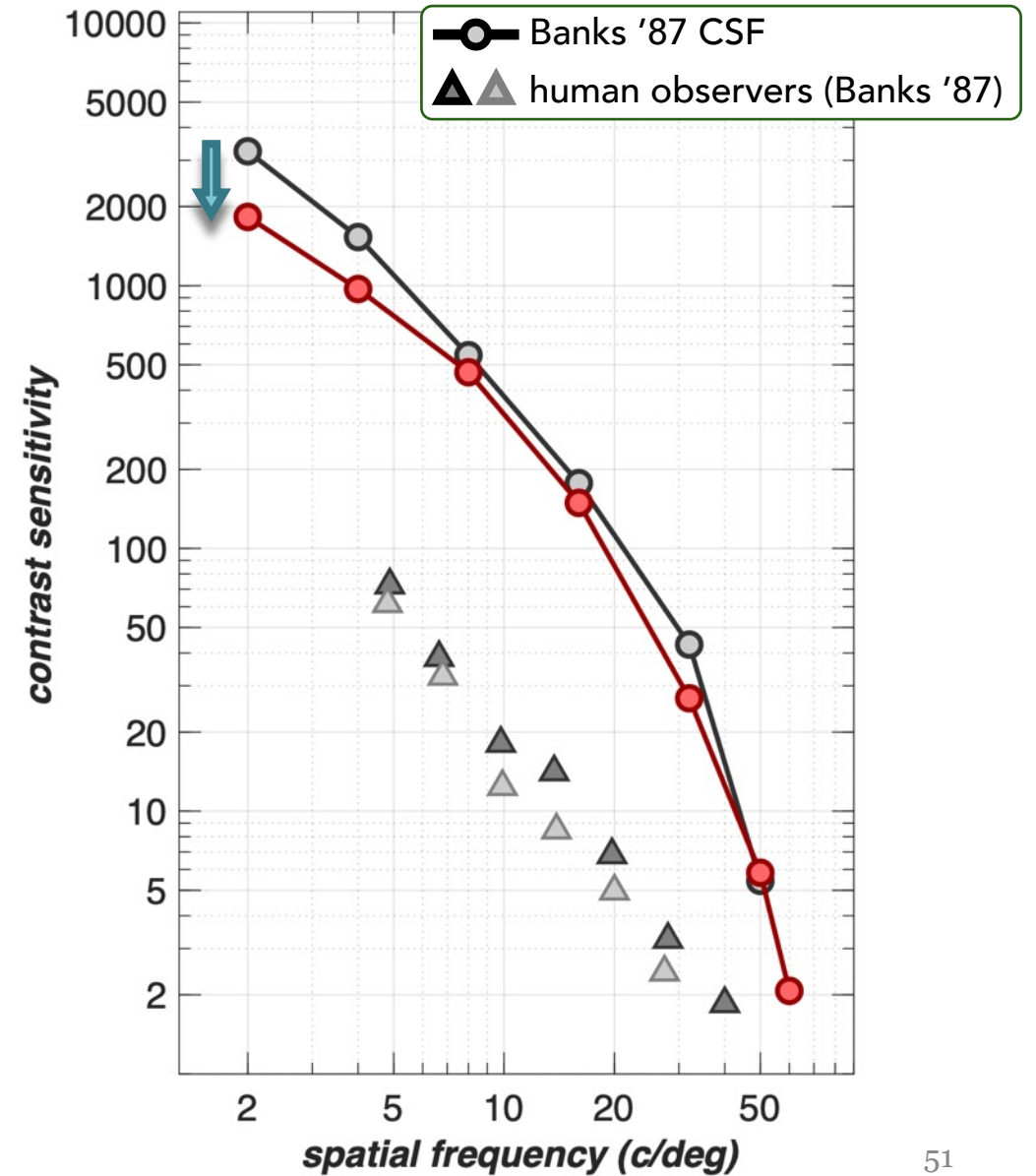


(from Thibos et al. '92)

Eccentricity-dependent density mosaics

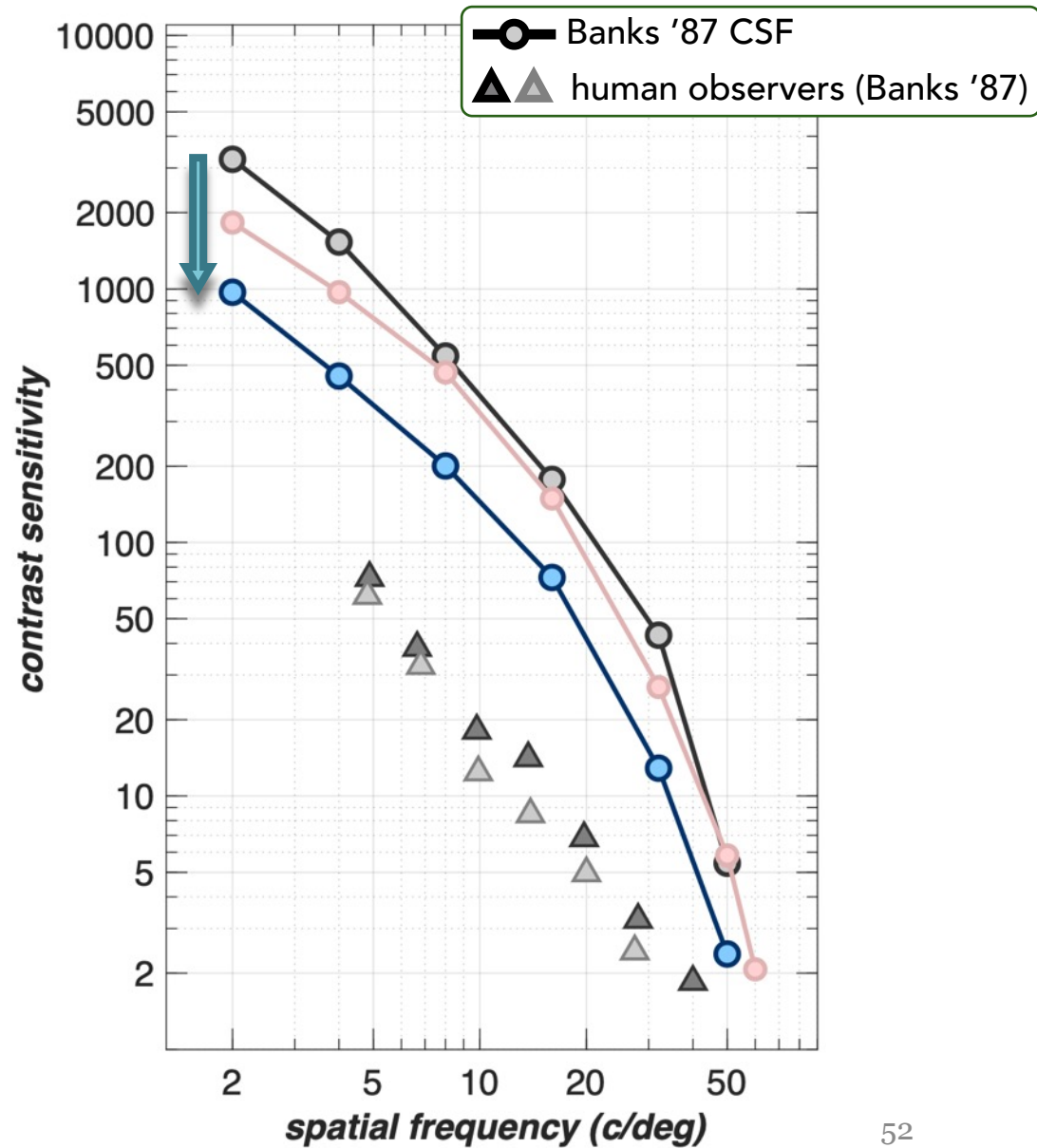
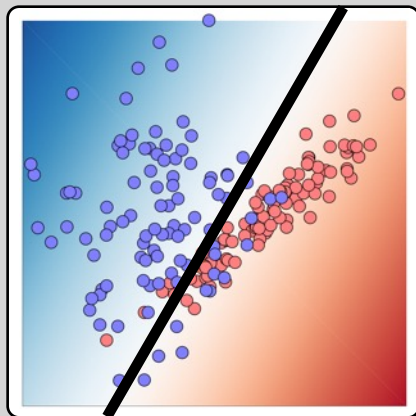
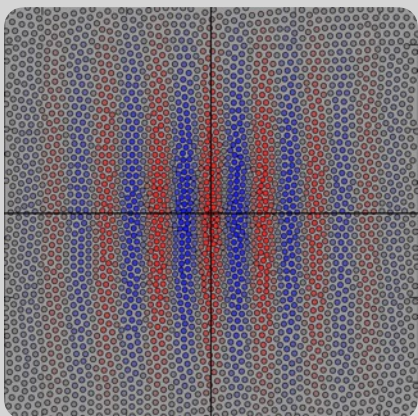
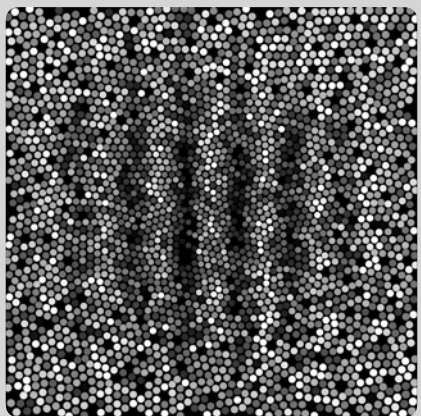


(based on Curcio et al. '90)

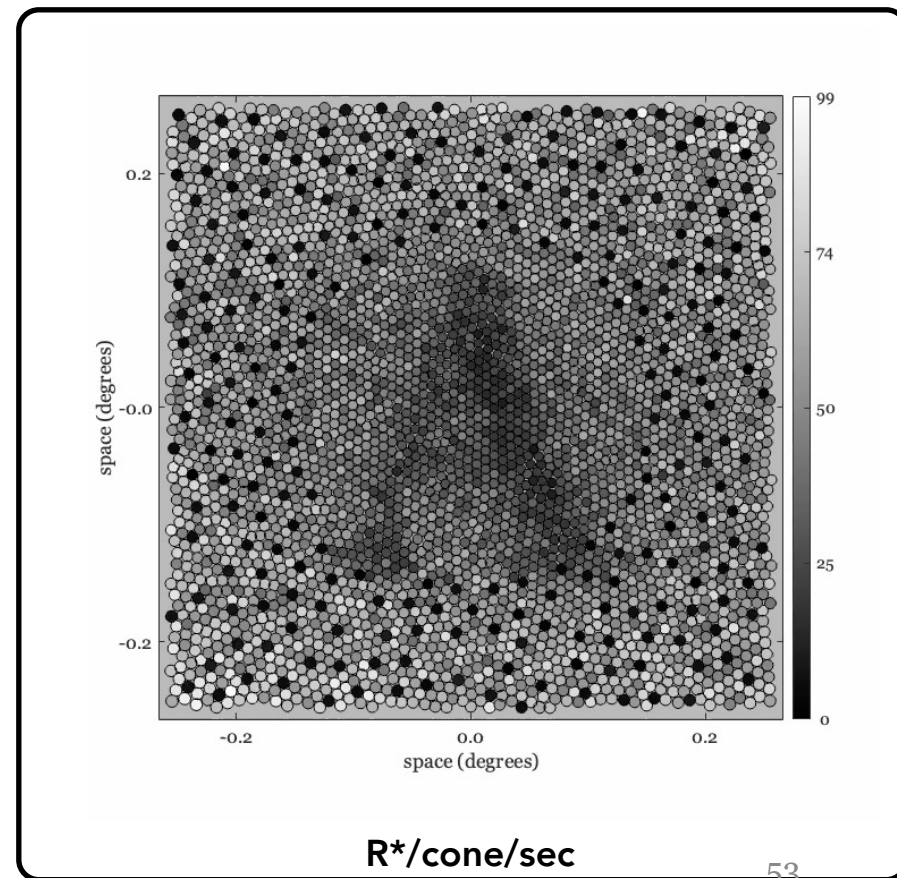
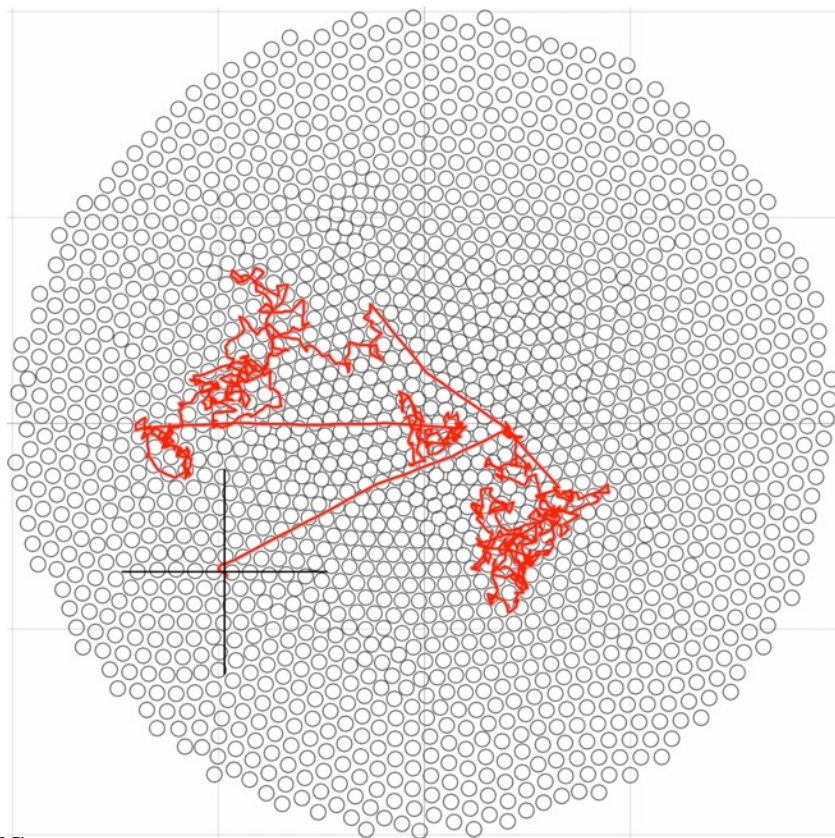
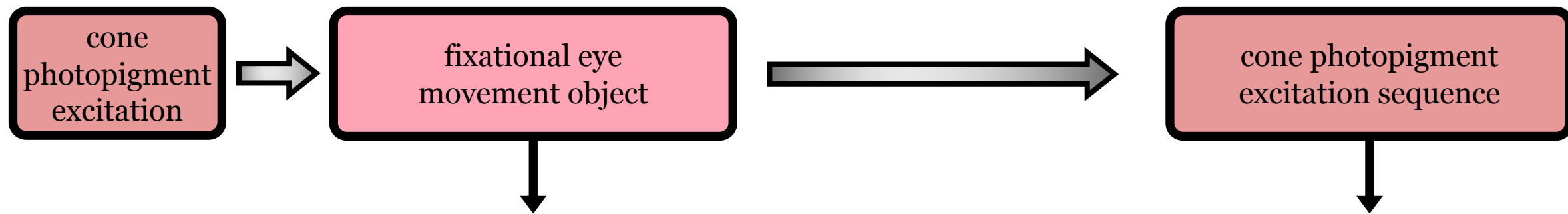


Accounting for absolute sensitivity: partially learned classifier

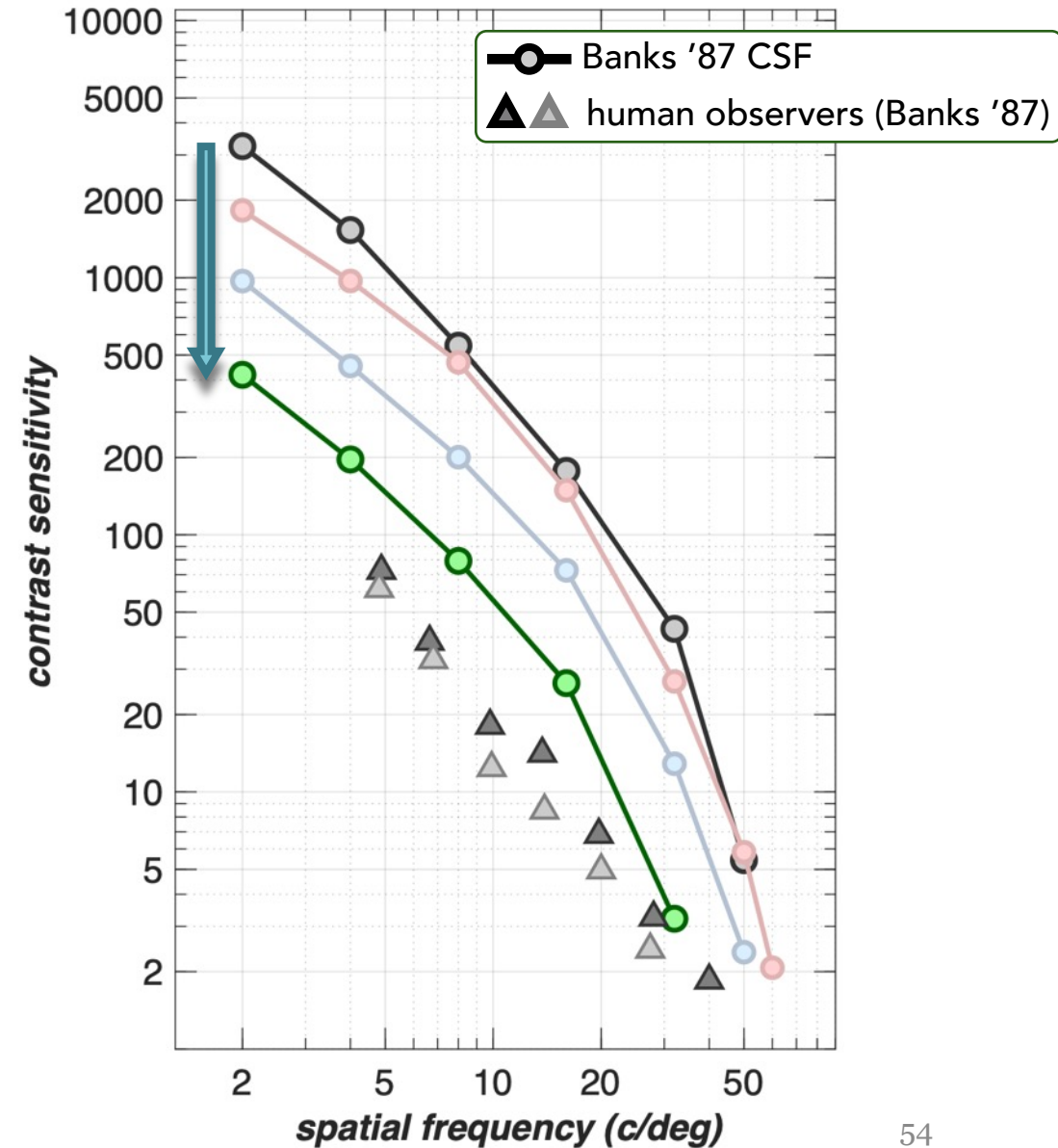
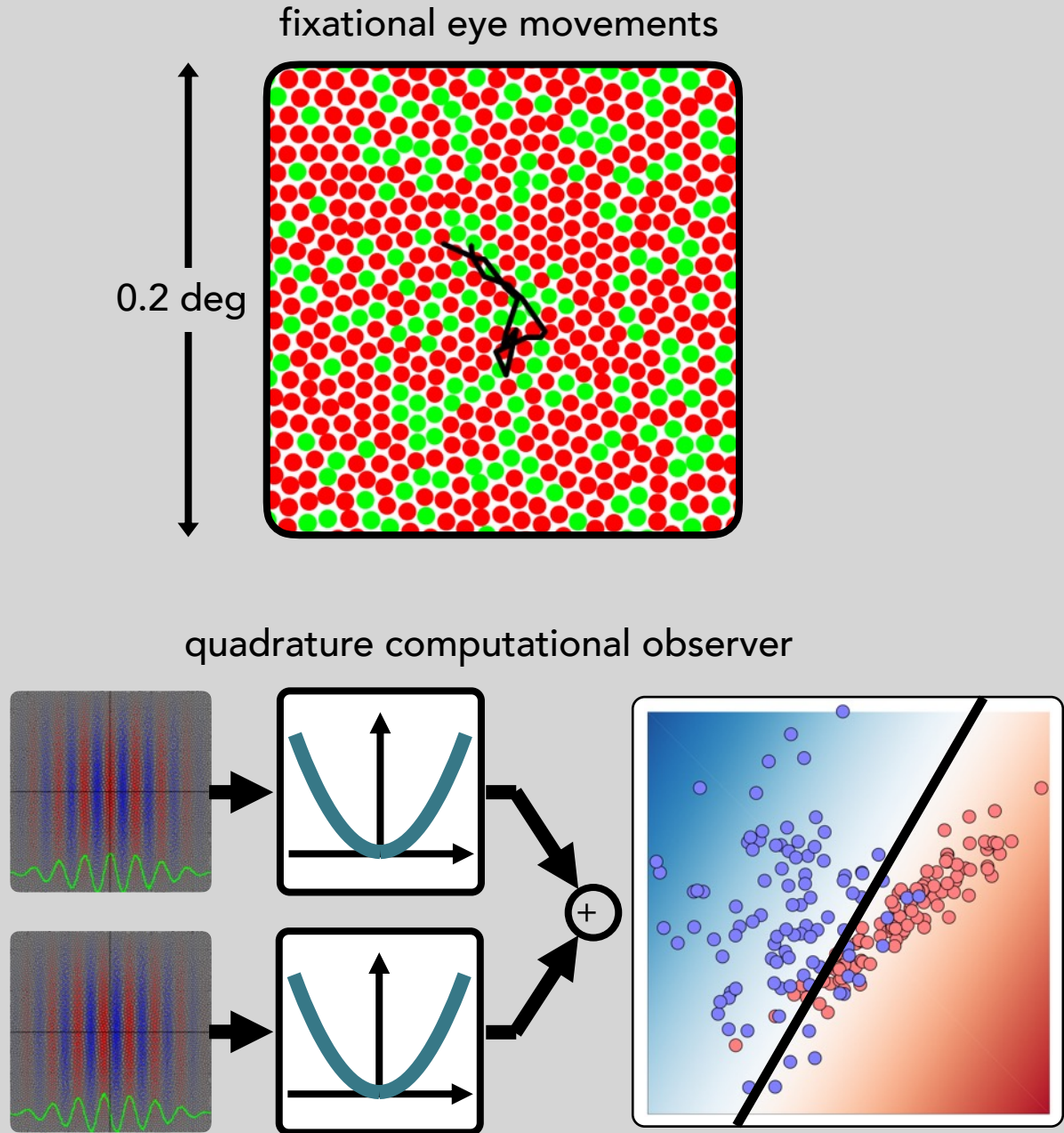
Computational observer



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

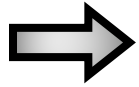


Energy computations (quadrature filters) reduce the impact of fixational drift

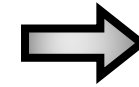


Time: Photocurrent transduction (Angueyra & Rieke, 2013)

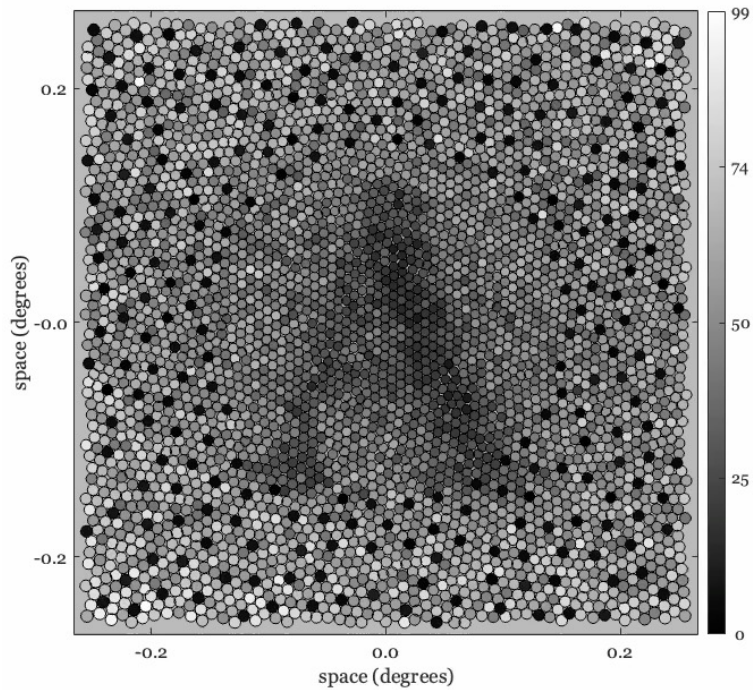
cone photopigment
excitation sequence



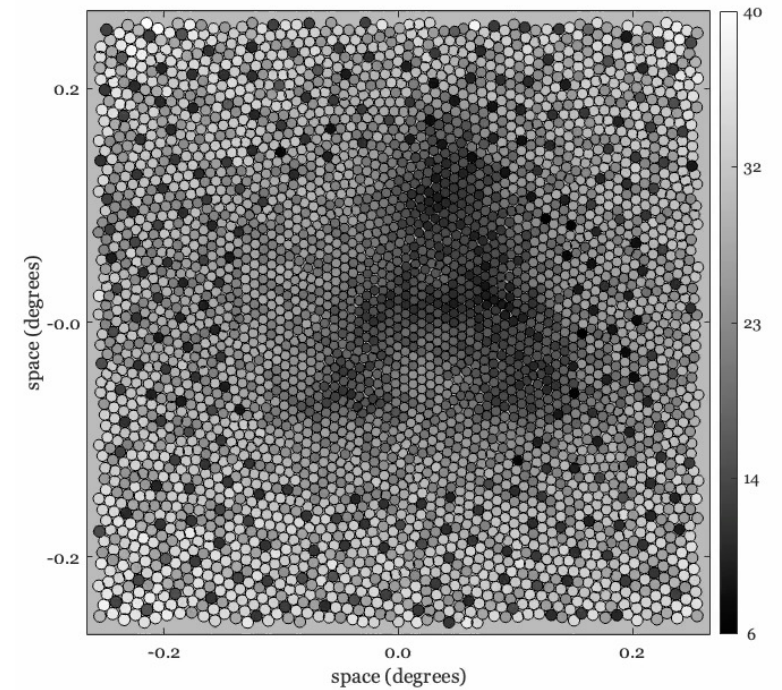
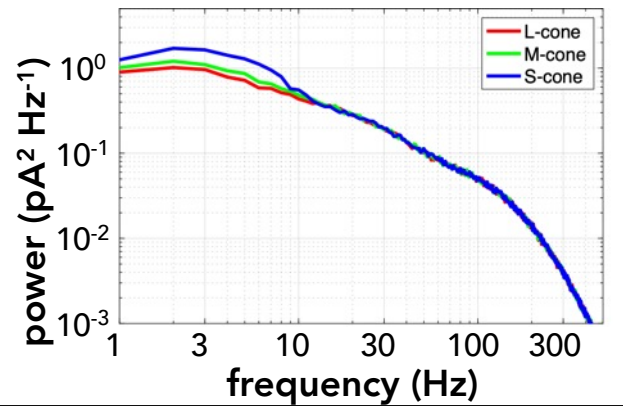
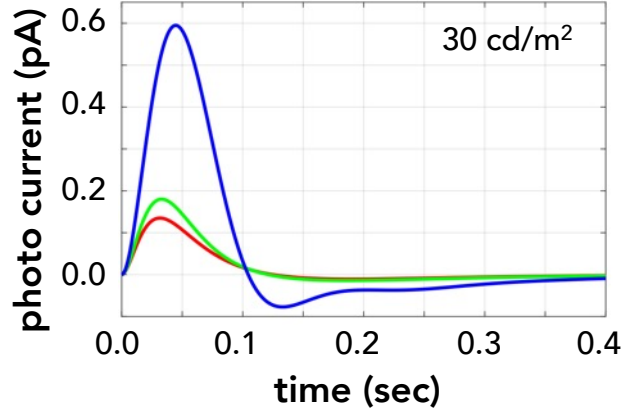
outer segment
object



cone outer-segment
photocurrent response



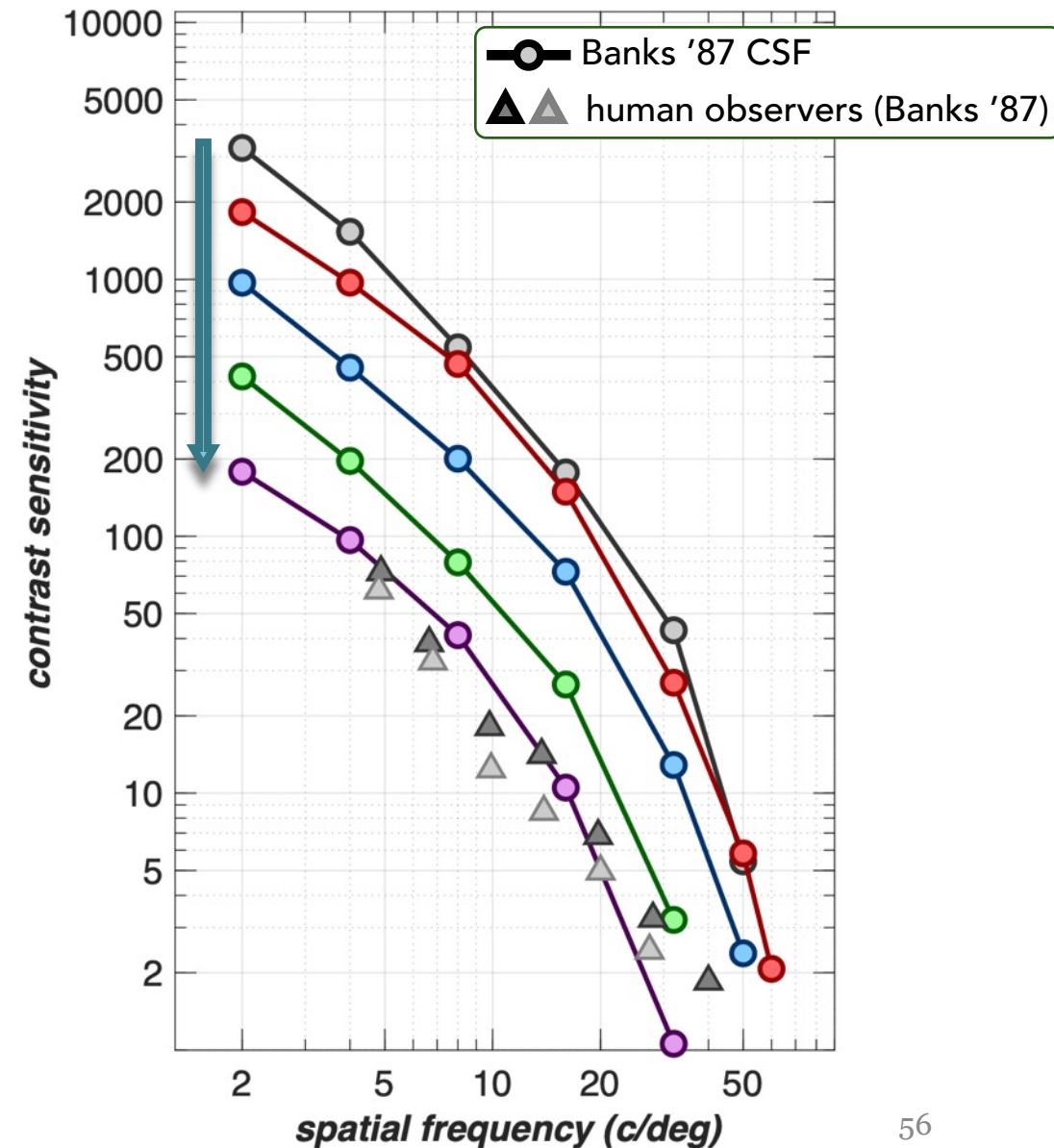
$R^*/\text{cone}/\text{sec}$



pAmps

Accounting for absolute sensitivity: photocurrent transduction

1. **Updated optics & cone mosaic** modeling has a minor impact relative to the Banks '87 estimate (factor of 1.7 at 2 c/deg),
2. **Computational observers**, which learn visual tasks by observing neural responses, result in a significant sensitivity drop across the entire spatial frequency range (accumulated factor of 2-3).
3. Inclusion of **fixational eye movements**, requires non-linear computational observers, and further reduces sensitivity across the entire spatial frequency range (accumulated factor: 7-10).
4. Inclusion of **photocurrent encoding** further reduces sensitivity approaching psychophysical limits (accumulated factor: 18-30).

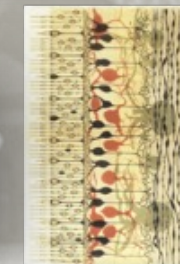
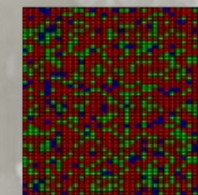
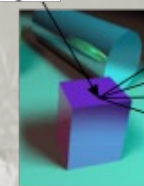


Third point

Modeling the the visual pathway components can be helpful in clarifying how the complex set of biological factors combine to limit human performance.

Image systems engineering tools (ISET): Starting from 2018

- **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)
- **ISETBio** is a Matlab toolbox that computes the visual encoding



Simulation technologies for image systems engineering and biology

Brian A. Wandell

Stanford Center for Image Systems Engineering

Wu Tsai Neurosciences Institute
Stanford Center for Cognitive and
Neurobiological Imaging

QUANTITATIVE MEASUREMENTS

∞

COMPUTATIONAL MODELS

∞

CHECK AND SHARE

