

# 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

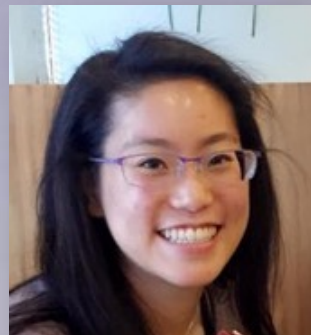
QUANTITATIVE MEASUREMENTS

∞

COMPUTATIONAL MODELS

∞

CHECK AND SHARE

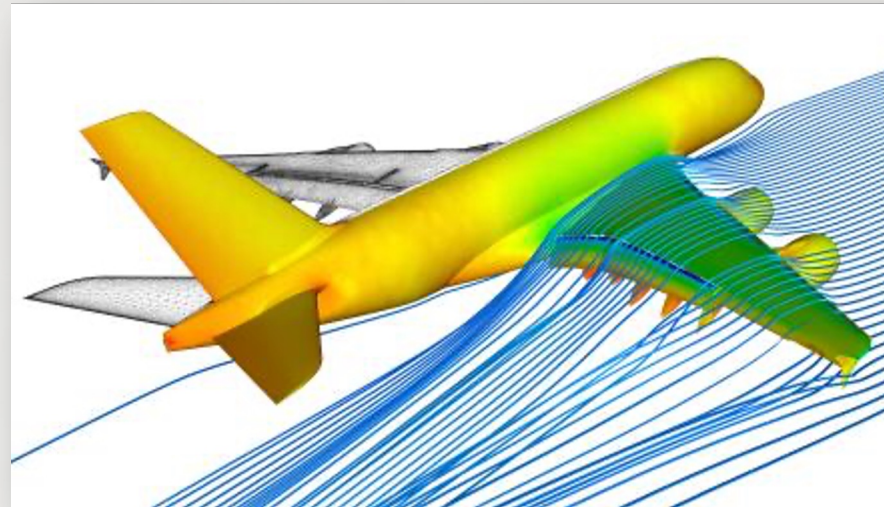


# Systems simulation is important in many mature industries

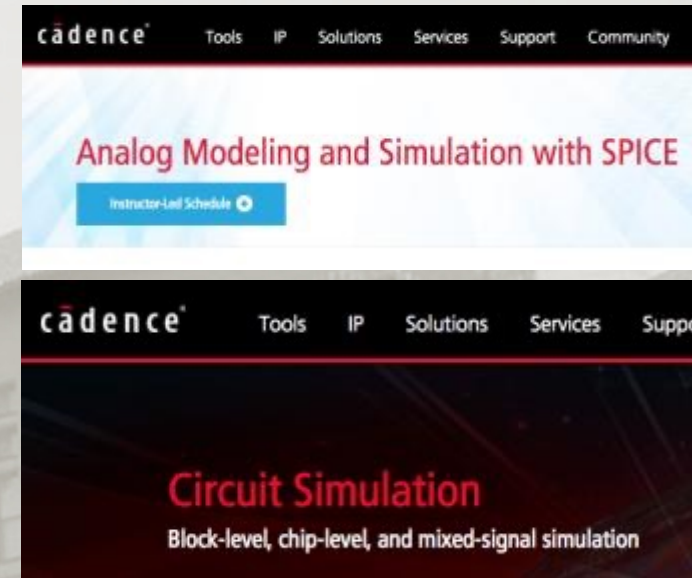
Image systems have become mature



**ECU (Electronic Control Unit) Simulation for Automobiles**



**Numerical flow simulation on an Airbus A380**



**Integrated circuitry**

# 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

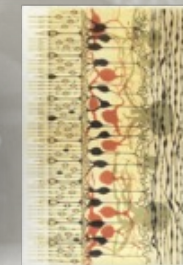
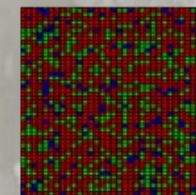
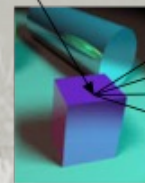
Open Sourced on GitHub  
in 2018



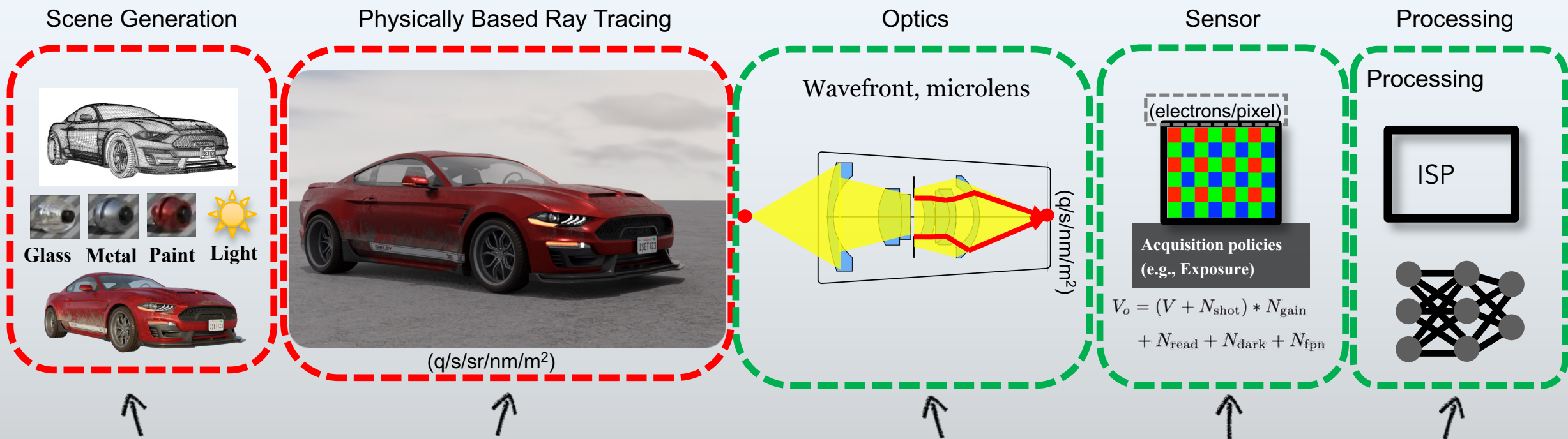
Image systems simulation software that is trusted by key stakeholders in industry and academia can speed the development of next generation image sensors, camera arrays and displays.

# Image systems engineering toolbox (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)
- **ISETBio** is a Matlab toolbox that computes the visual encoding, including lens and photoreceptor absorptions



# Simulation pipeline: ISETCam



ISET3d

ISETCam

Calculate the spectral irradiance at the sensor of complex scenes.

Calculate sensor response to the spectral irradiance; Many tools for evaluation of color (CIE standards) and spatial resolution (ISO standards).

# Simulation pipeline: ISETCam-ISETBio

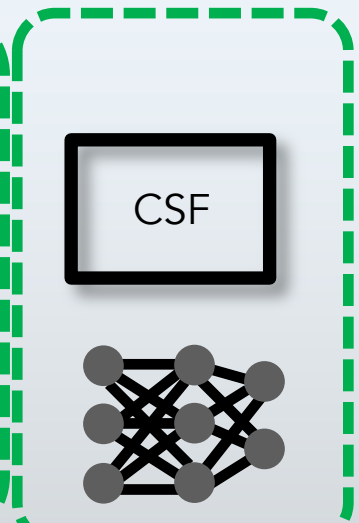
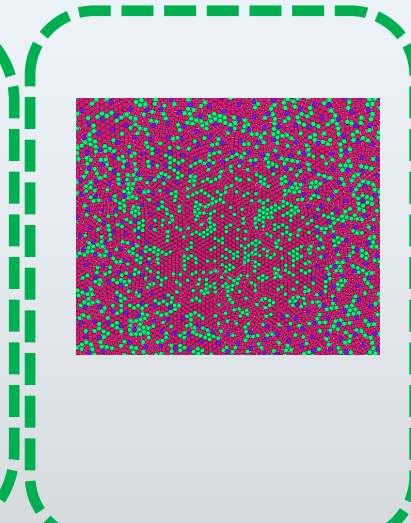
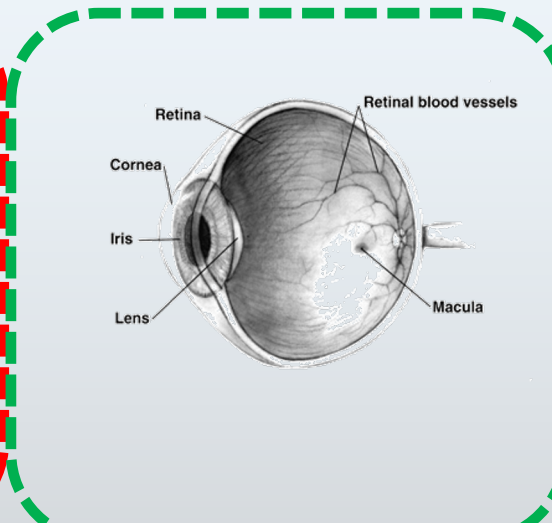
Scene Generation

Physically Based Ray Tracing

Human optics

Retinal modeling

Processing



ISET3d

ISETCam-ISETBio

Calculate the spectral irradiance at the sensor of complex scenes.

Calculate retinal response to the spectral irradiance;  
Also, many tools for modeling and visualizing biological processes

# Software organization and validation

Unit, system  
regression tests

isetvalidate  
UnitTestFixture  
ExampleTestFixture

Library routines  
tutorials  
examples

ISETCam  
ISET3d

ISETBio

Device projects

ISETAuto ISETRtf ISETLens  
ISETCornellbox ISETFluorescence  
ISETFundamentals

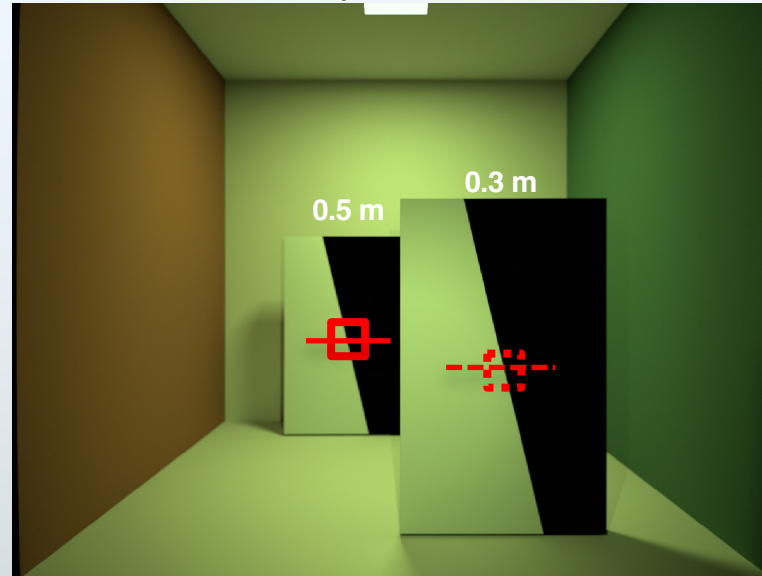
Biology projects

ISETBioCSFGenerator  
ISETBioPrinciples  
ISETBioAppearance

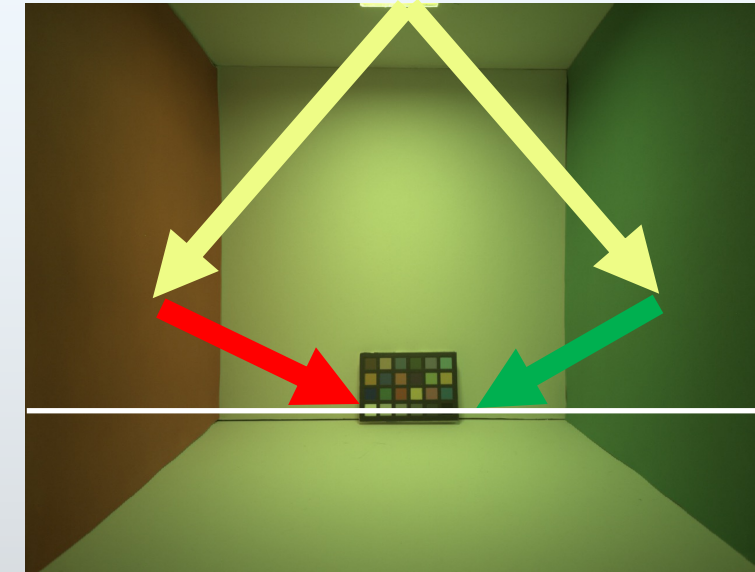
# Experimental validation

- Experimental validation of the software, in addition to software validation

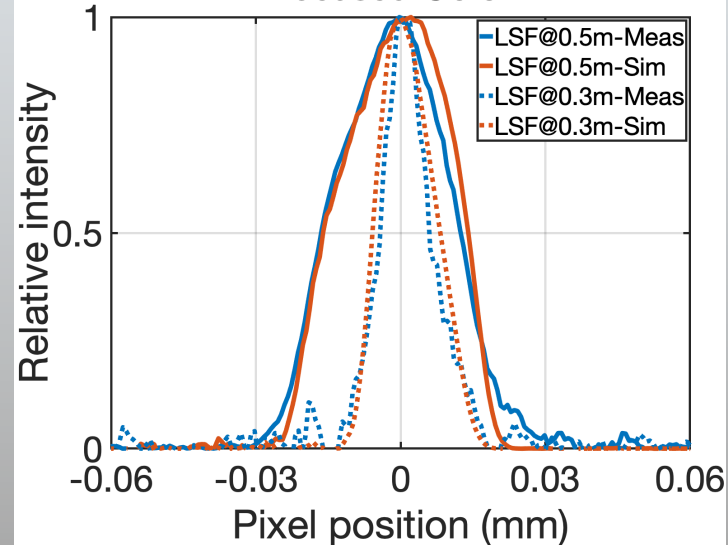
Resolution and Depth of Field



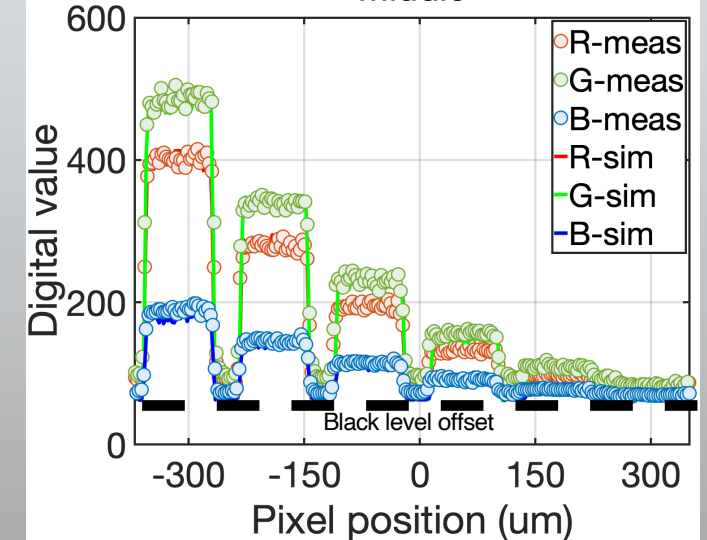
Inter-reflections and electronic sensor noise



Focused @0.3m

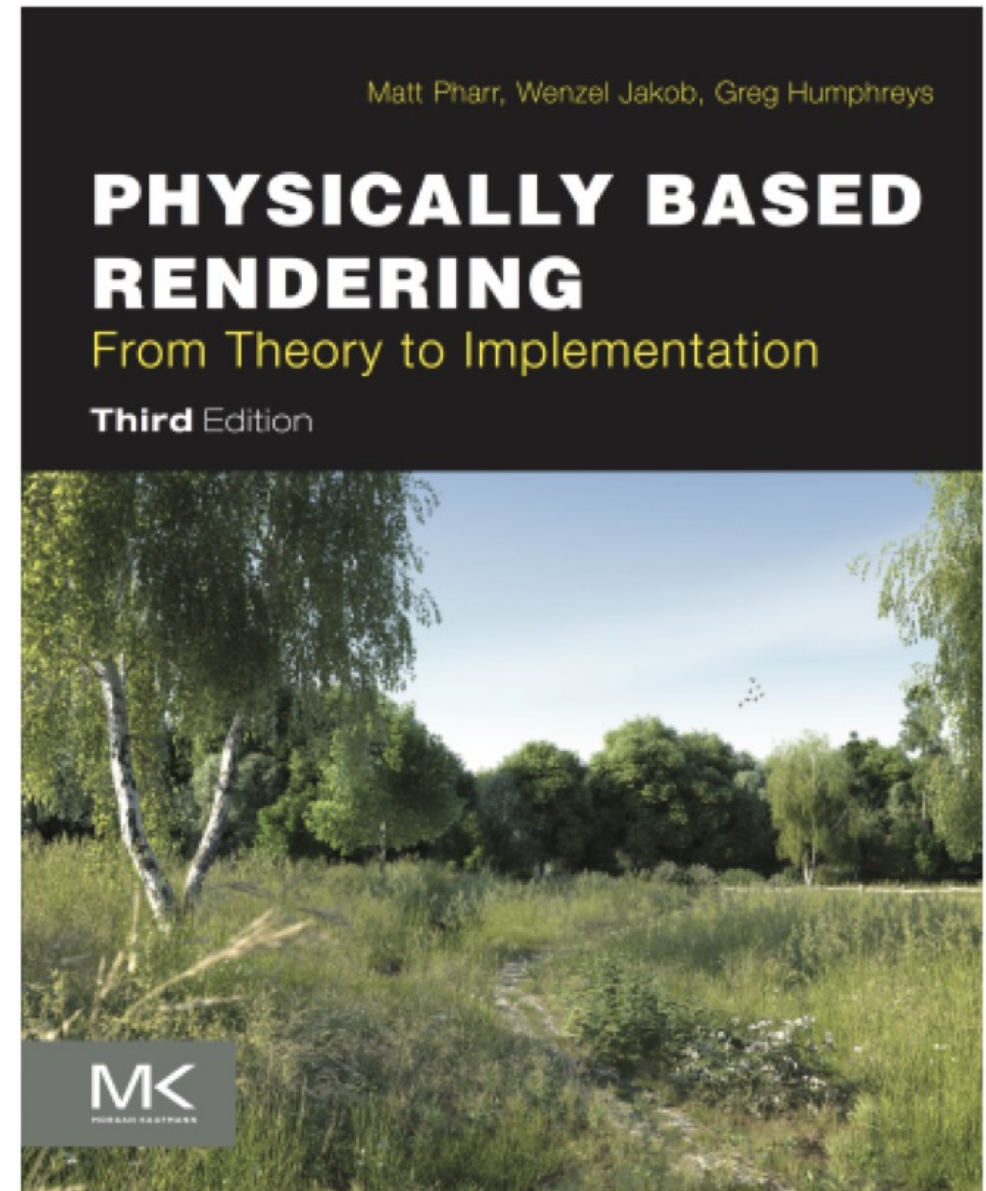


Middle



# Quantitative computer graphics is necessary component for materials and lights

- 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!

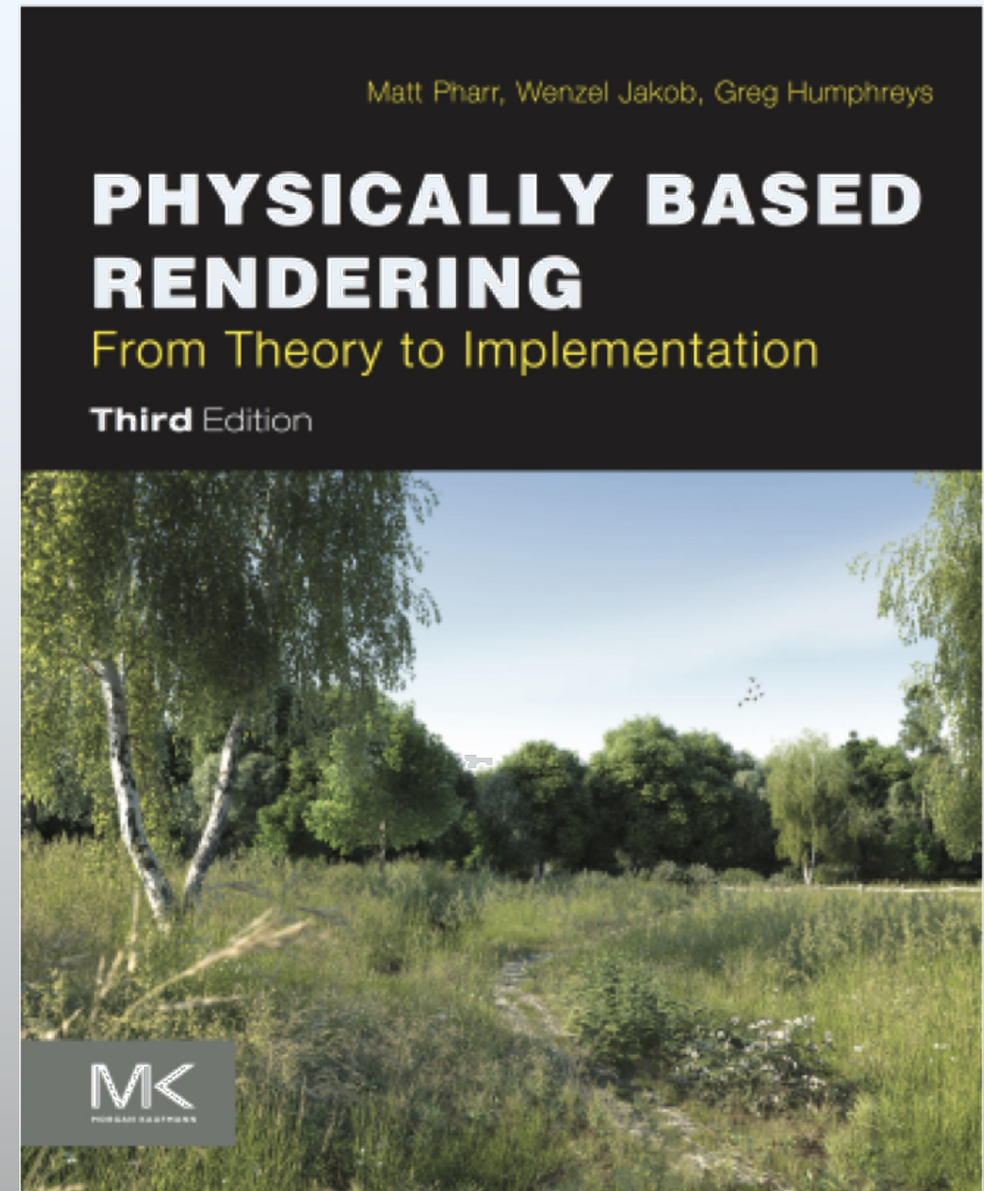


# Quantitative computer graphics is necessary component for materials and lights

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)



# Standard optics

Sensor: 7.5mm diagonal

Double Gauss



FOV: 22 deg  
Focal length: 3 mm

Wide angle

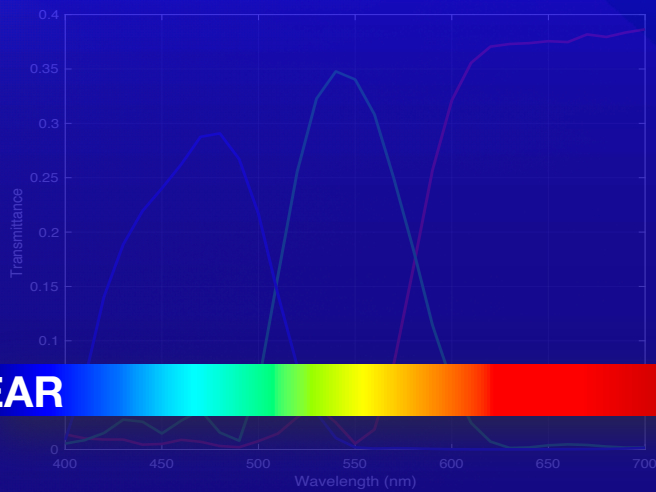
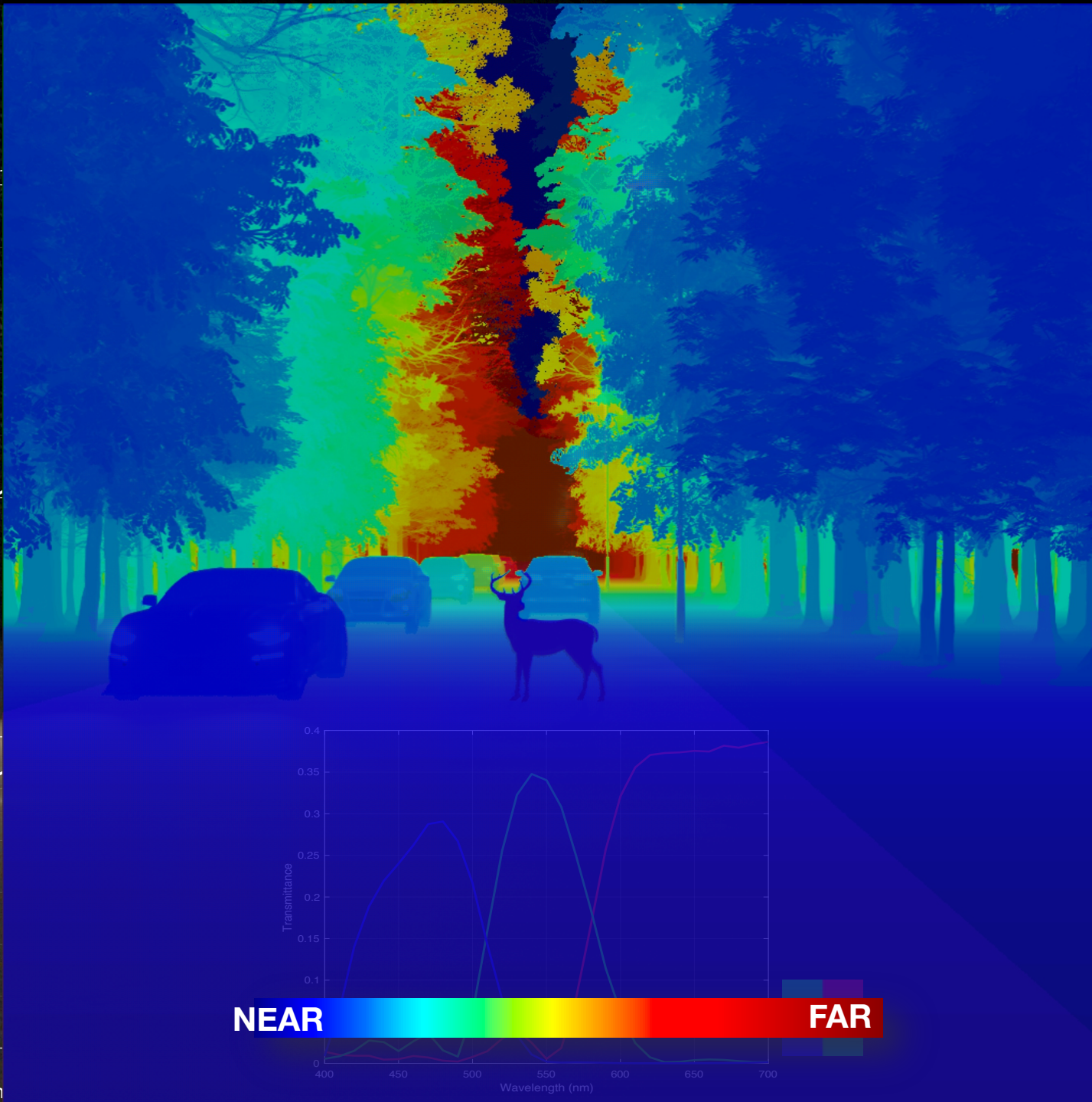


FOV: 56 deg  
Focal length: 3 mm

Fisheye



FOV: 87 deg  
Focal length: 3 mm

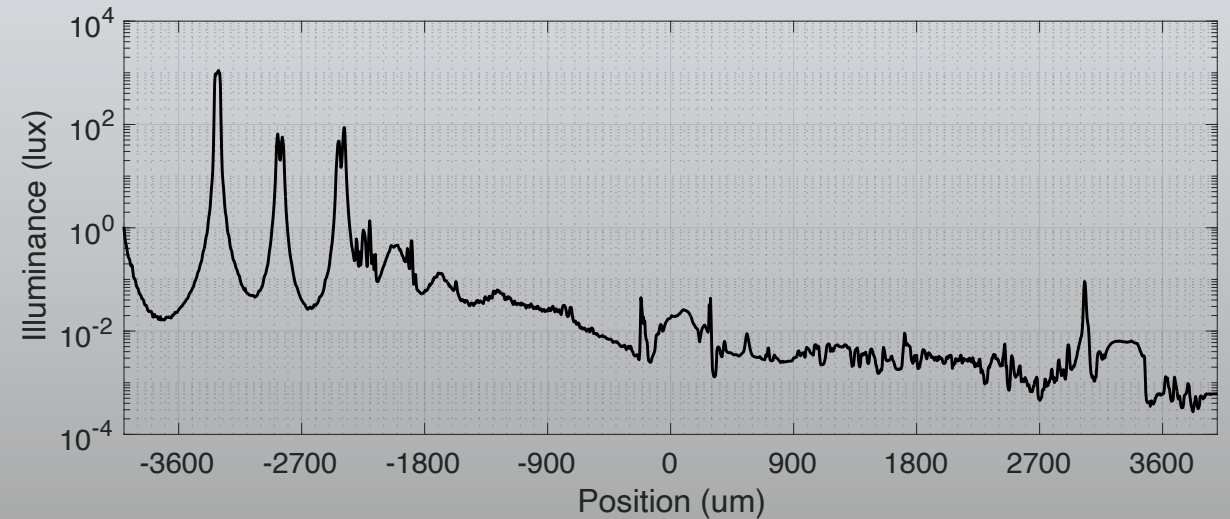
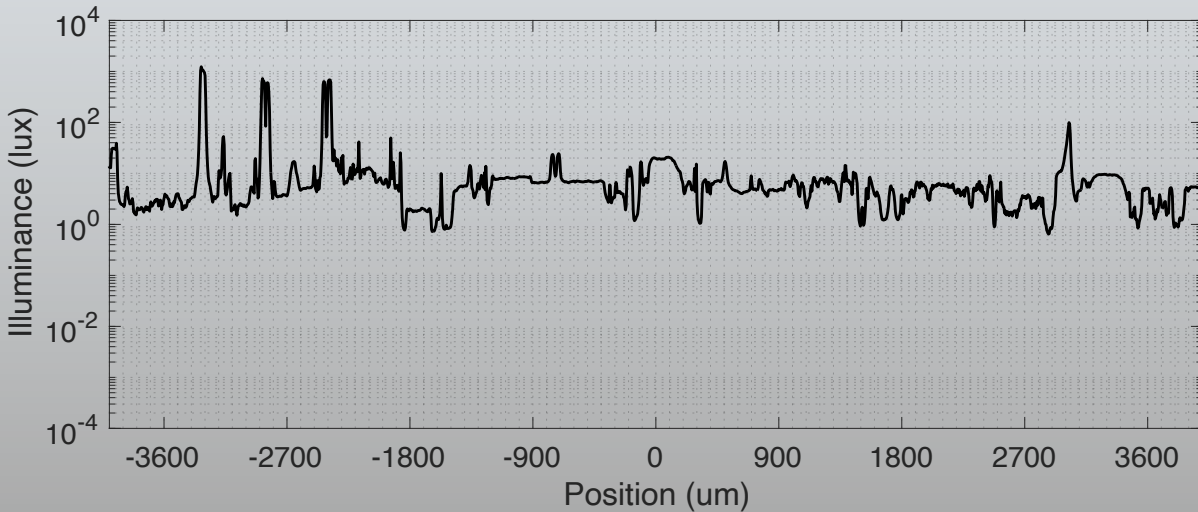


# RGB images – representing high dynamic range data

DR: 1000:1



DR: 1,000,000:1



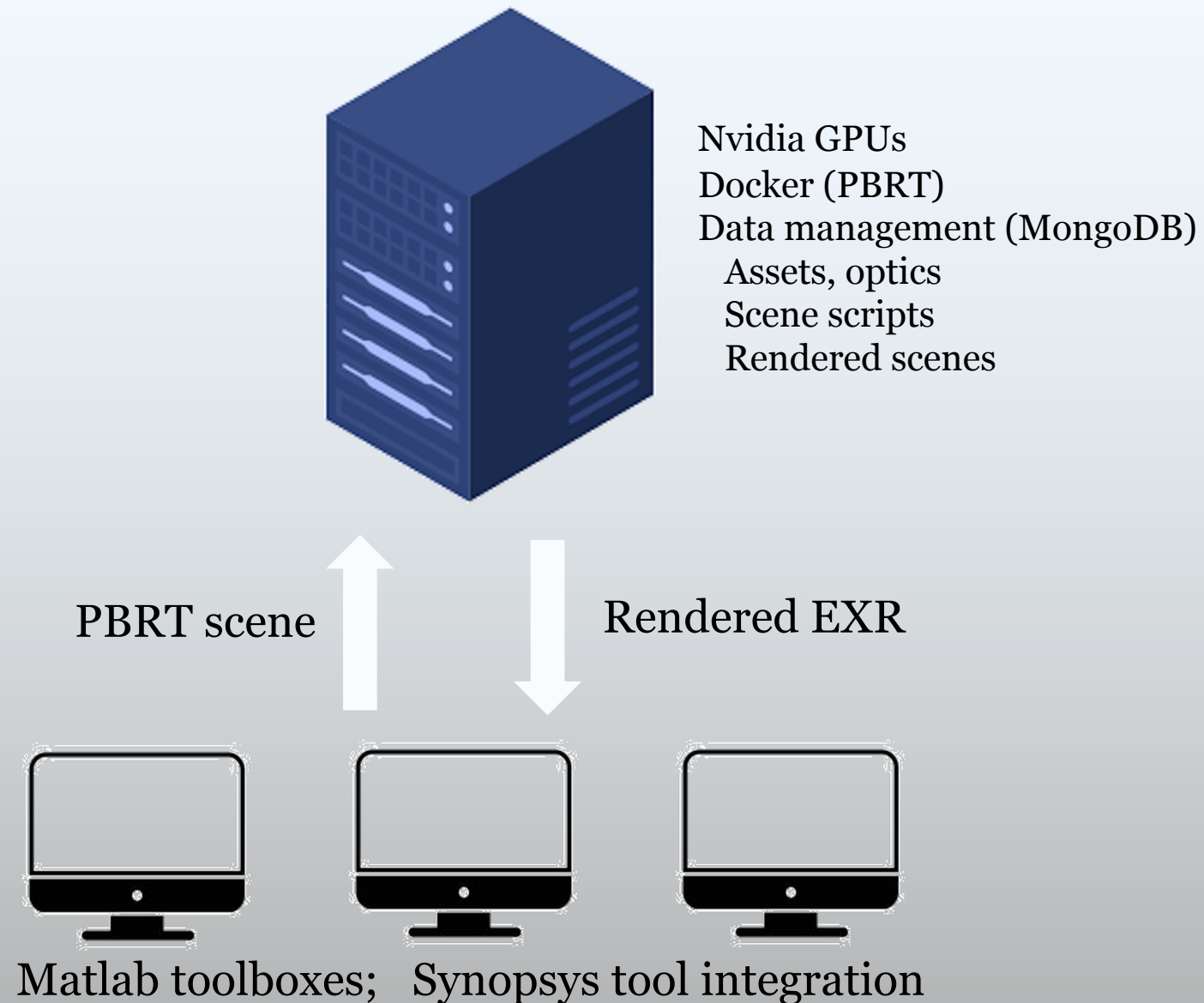
# Data and compute values: Single scene

- Imaging optics and microlens data
- Image sensors from 3M to much larger
- Data management of assets, scenes (Terabytes)
- Compute using GPUs and parallelized through Dockers/Kubernetes
- Time: 400 hours for synthesis of a modest training set; a few hours for fine-tuning

Category	Parameter	Value
Spectral Radiance	Number of pixels	1920 x 1080 (3M)
	Rays per pixel	2000 rays per pixel
	Wavelengths	31 wavebands
	Memory	20GB
Assets	Objects	1,000
	Polygons per asset	100,000
Single Image Rendering Time (High-End GPUs)	CPU	60 minutes
	GPU (Nvidia 3090 or 4090)	5 minutes
Number of Images Rendered	Network Fine-tuning	5,000
	Training from Scratch	» 5,000

# Distributed computing infrastructure

- Imaging optics and microlens data
- Image sensors from 3M to much larger
- Data management of assets, scenes (Terabytes)
- Compute using GPUs and parallelized through Dockers/Kubernetes
- Time: 400 hours for synthesis of a modest training set; a few hours for fine-tuning





Zhenyi Liu

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

Zhenyi Liu<sup>1,2</sup>, Minghao Shen<sup>2</sup>, Jiaqi Zhang<sup>3</sup>, Shuangting Liu<sup>3</sup>, Henryk Blasinski<sup>1</sup>, Trisha Lian<sup>1</sup>, Brian Wandell<sup>1</sup>  
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<sup>1,2</sup>, TRISHA LIAN<sup>2</sup>, JOYCE FARRELL<sup>2</sup>, AND BRIAN A. WANDELL<sup>2</sup>

<sup>1</sup>State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 13000, China

<sup>2</sup>Department 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

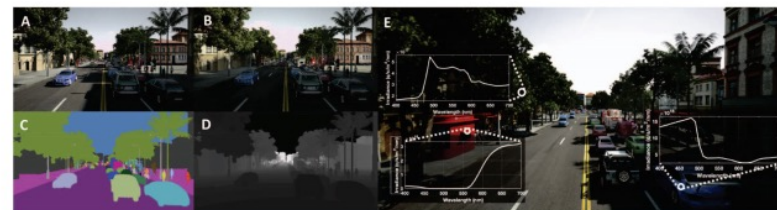


This ICCV Workshop paper is the Open Access version, provided by the Computer Vision Foundation. Except for this watermark, it is identical to the accepted version; the final published version of the proceedings is available on IEEE Xplore.

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

Zhenyi Liu<sup>1,2</sup>, Trisha Lian<sup>1</sup>, Joyce Farrell<sup>1</sup>, and Brian Wandell<sup>1</sup>

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



## ISETAuto: Detecting vehicles with depth and radiance information

ZHENYI LIU<sup>1</sup>, JOYCE FARRELL<sup>2</sup> AND BRIAN WANDELL<sup>2</sup>

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

<sup>2</sup>Stanford 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

# Companies responding to the need for synthetic data

**ANYVERSE**



**NVIDIA Omniverse**



**AirSim**



**Ansys**

**dSPACE**

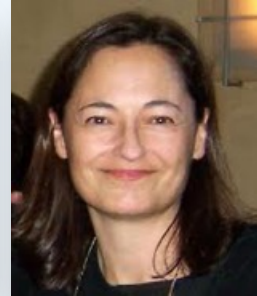


**datagen**

# What goes into an image systems simulation for the human eye



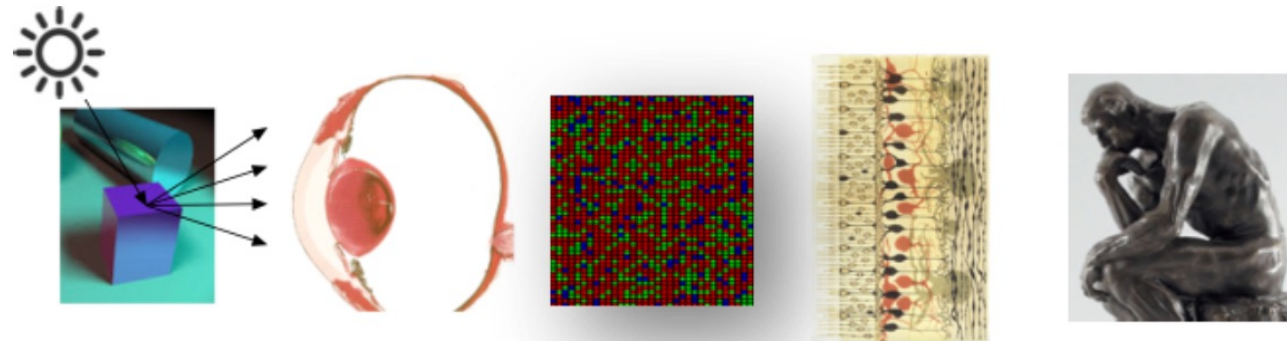
Nicolas Cottaris



Joyce Farrell



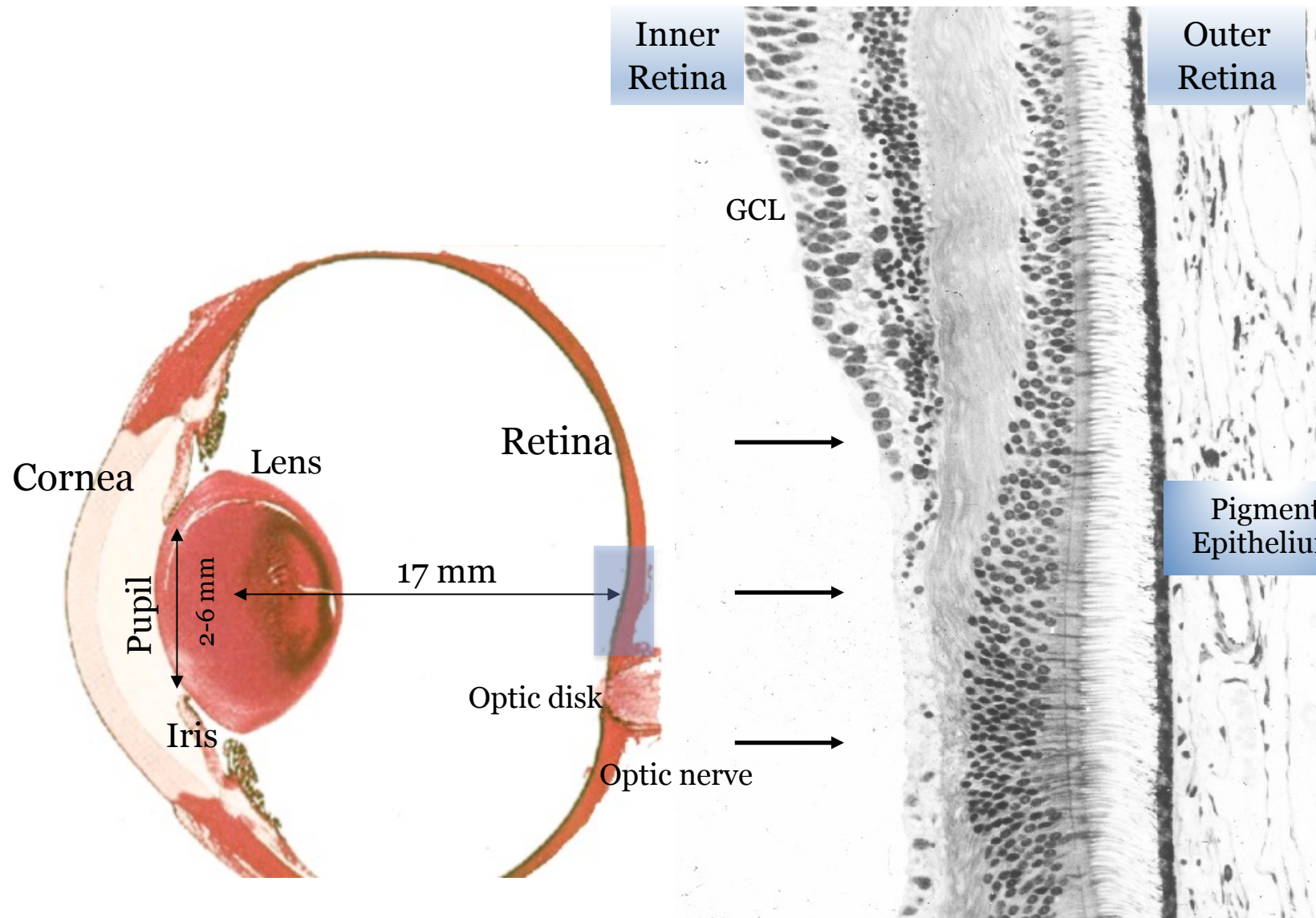
David Brainard



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

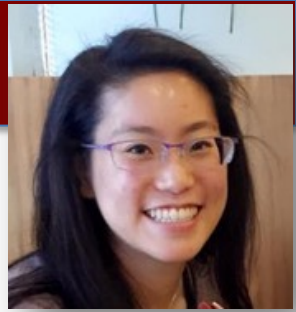
The Image System Engineering Toolbox for Biology (ISETBio) supports vision science calculations. The Matlab software enables you to create spectral radiance scenes and use these scenes as input to estimate the effects of human optics, eye movements, cone absorptions and photocurrent, and retinal cell 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
- $5 \times 10^6$  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

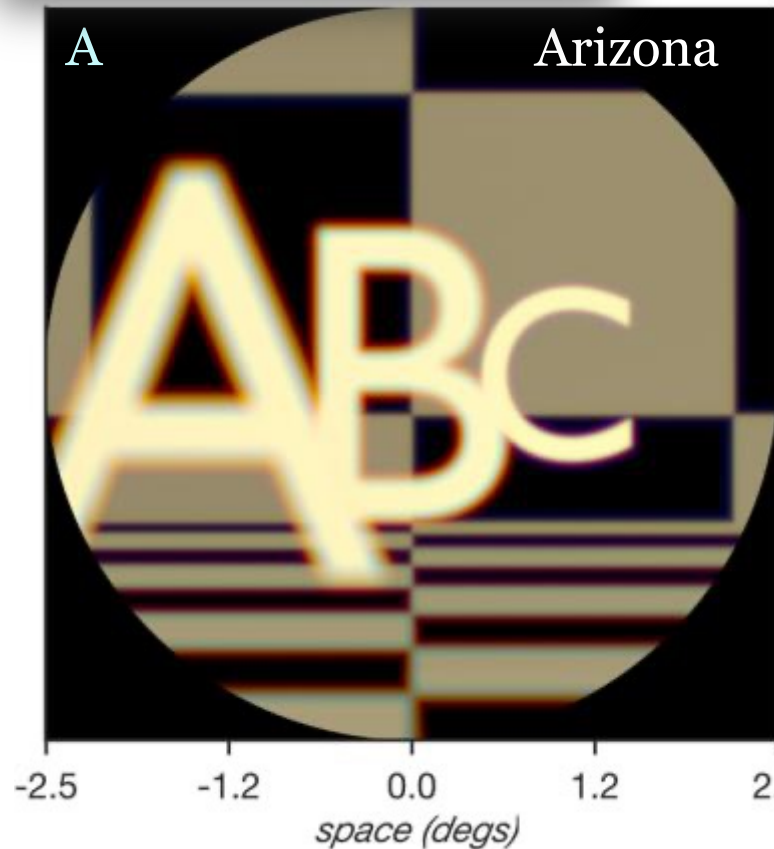


Trisha Lian

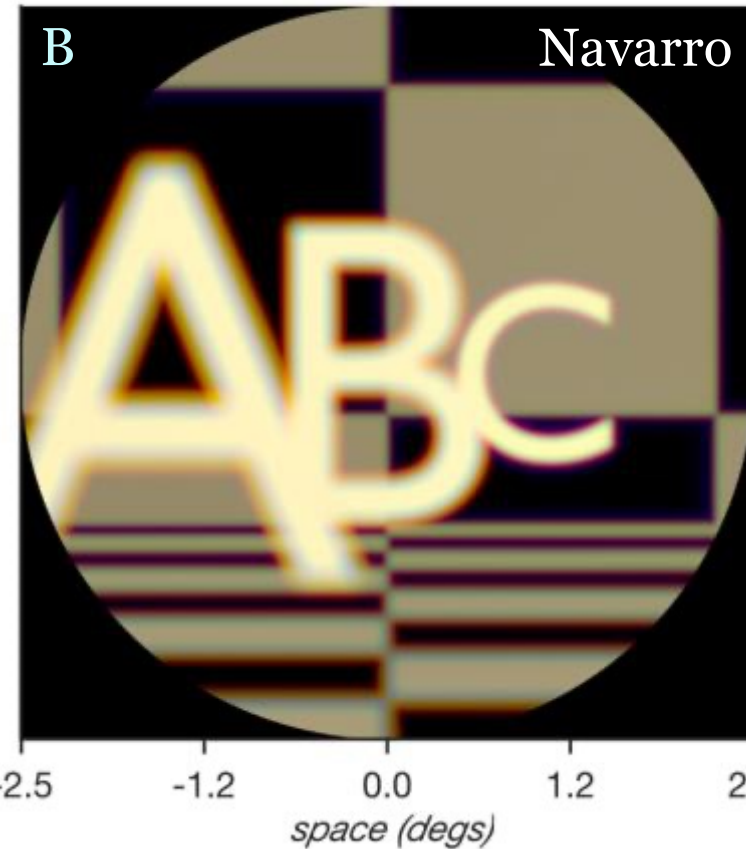
The PBRT-ISET3d code implements three major human eye models  
(Lian et al. 2019, Journal of Vision).

**Remember:** these images represent underlying spectral irradiance

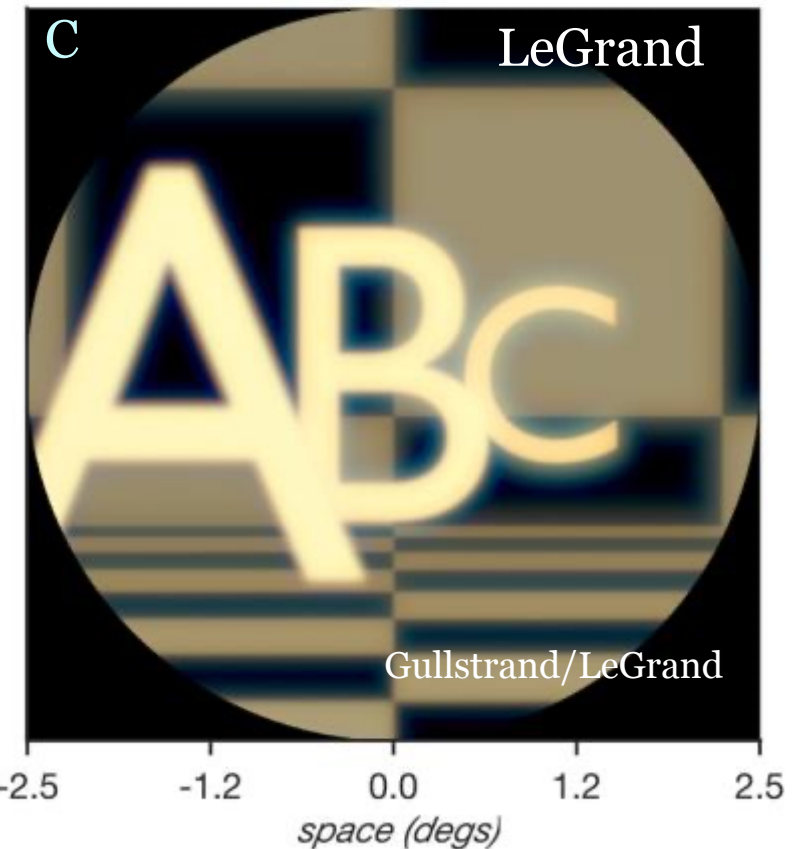
**OPEN ACCESS**  
Article | October 2019  
**Ray tracing 3D spectral scenes through human optics models**  
Trisha Lian; Kevin J. MacKenzie; David H. Brainard; Nicolas P. Cottaris; Brian A. Wandell  
[+ Author Affiliations](#)  
Journal of Vision October 2019, Vol.19, 23. doi:<https://doi.org/10.1167/19.12.23>



Schwiegerling, et al. 2004



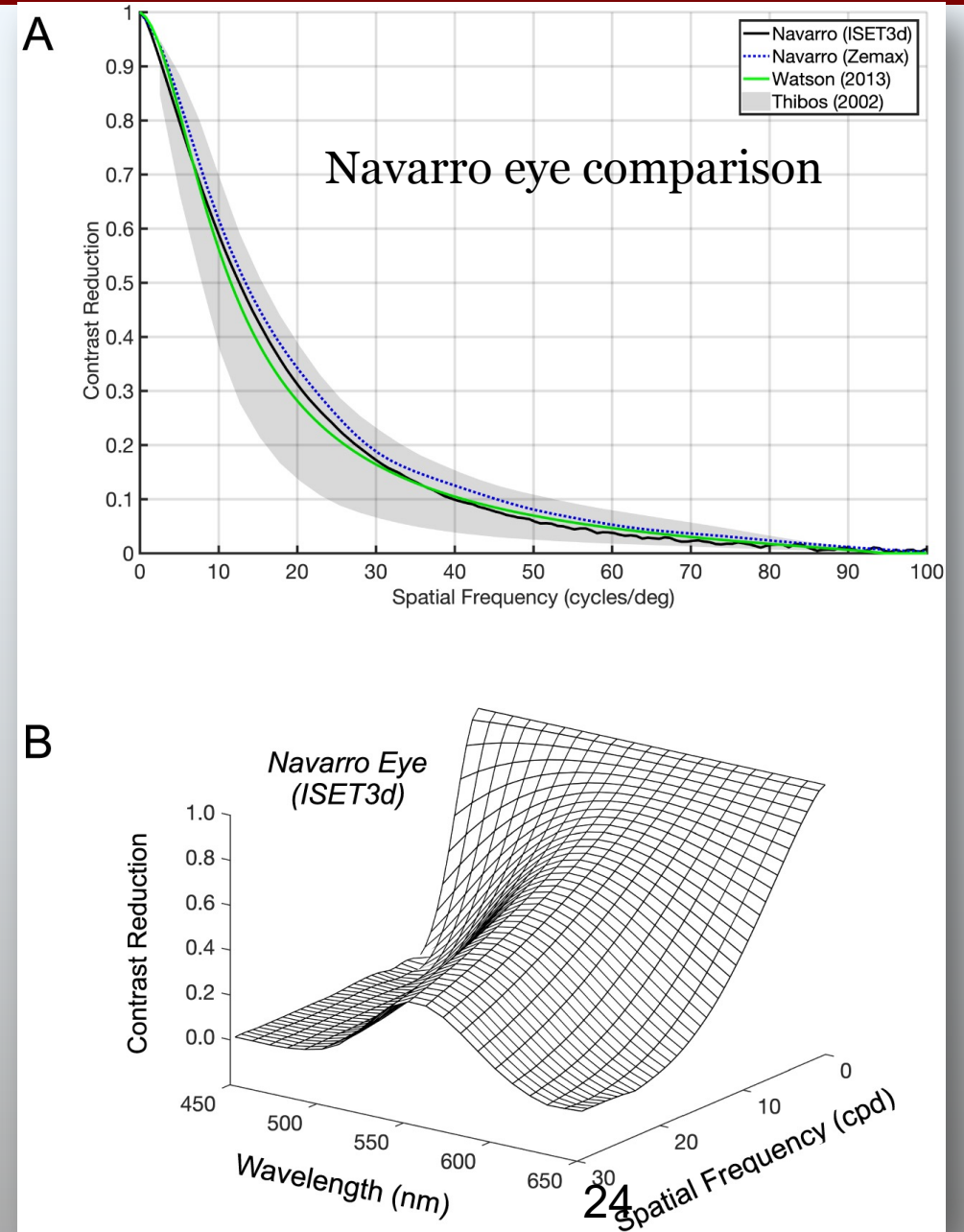
Escudero-Sanz & Navarro, 1999



Hage & Le Grand, 1980

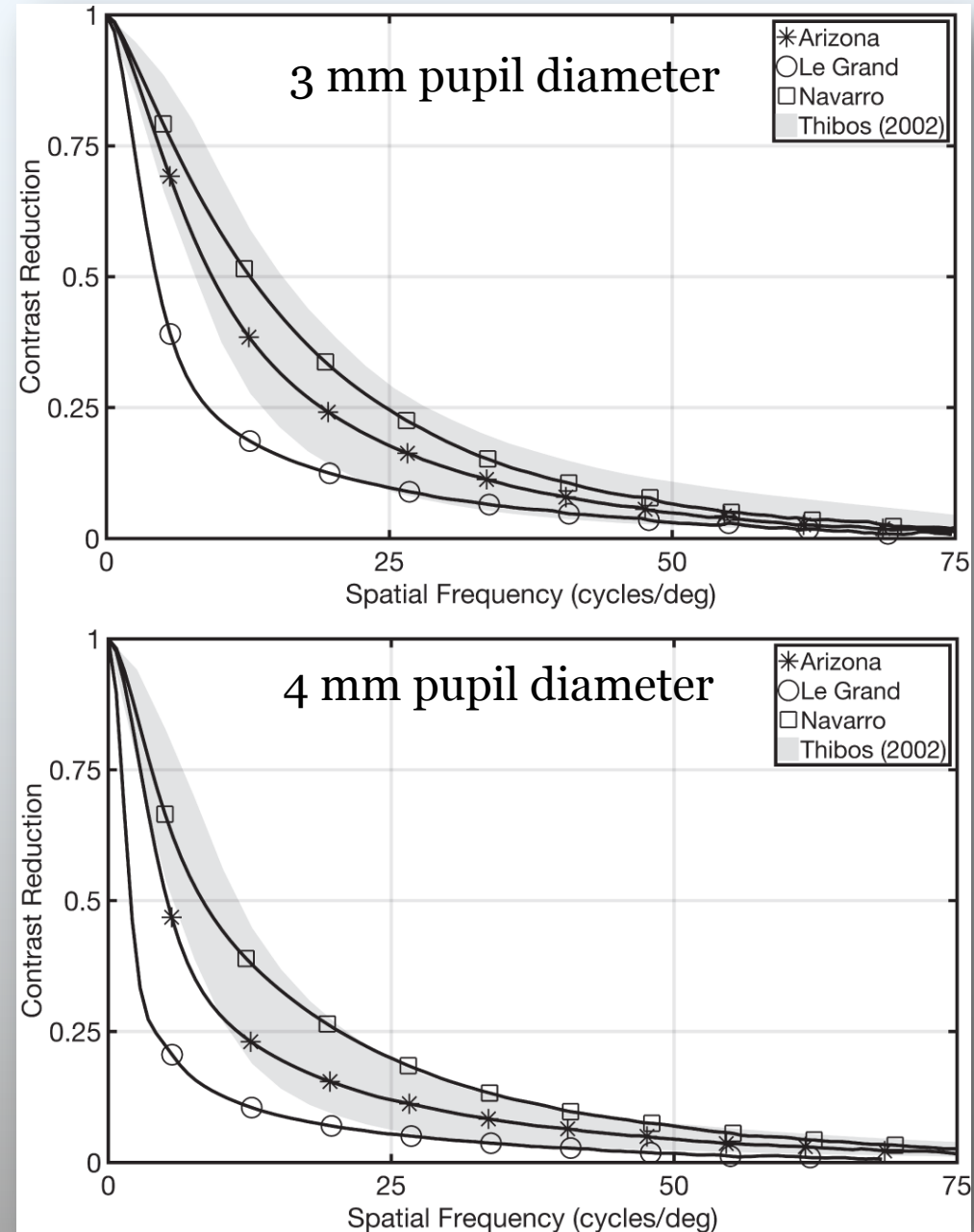
# Validation of the ray tracing methods

- 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



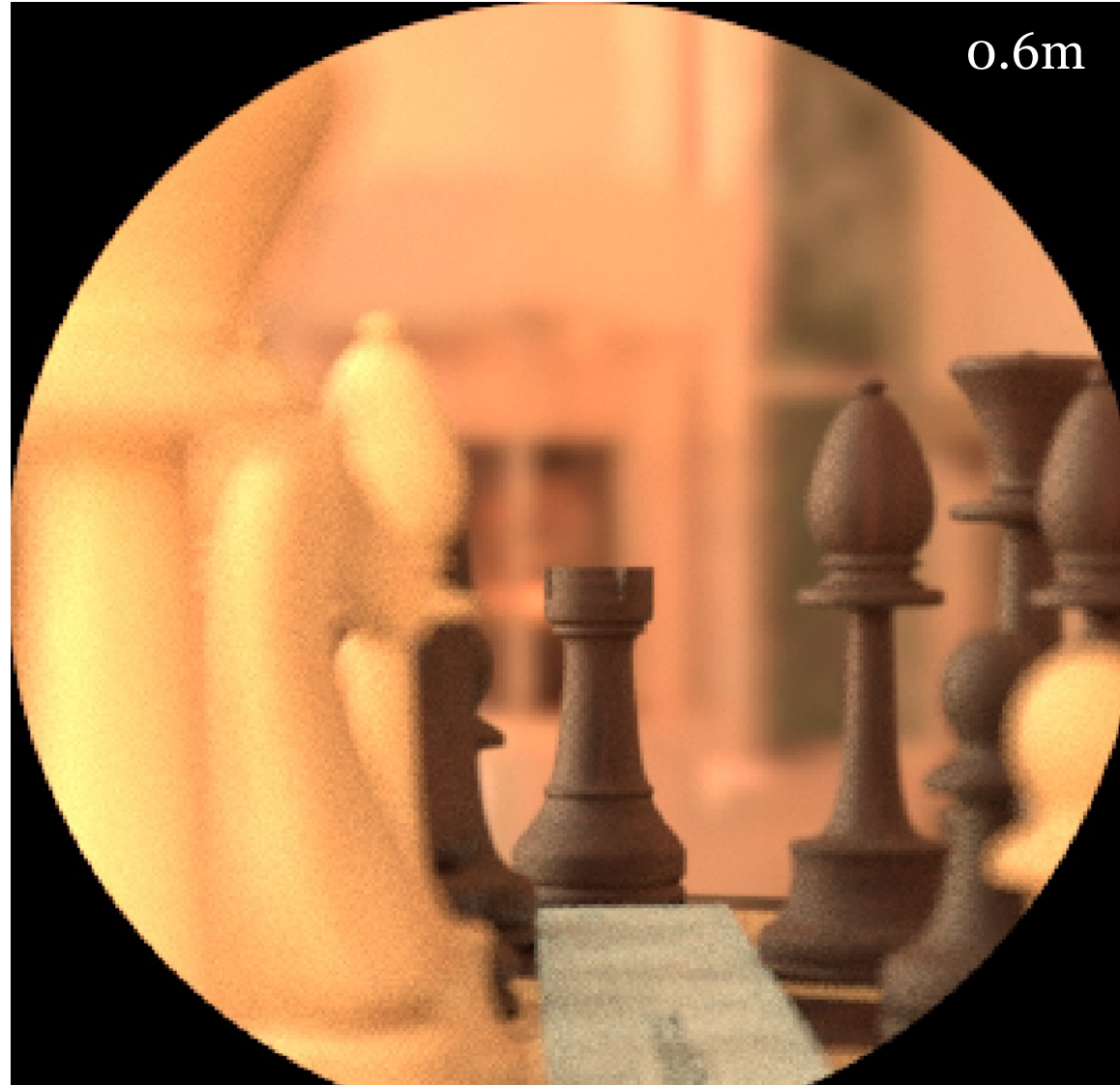
# Human eye model comparisons

- 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



## Depth of field and accommodation (5 mm pupil)

- The Navarro 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



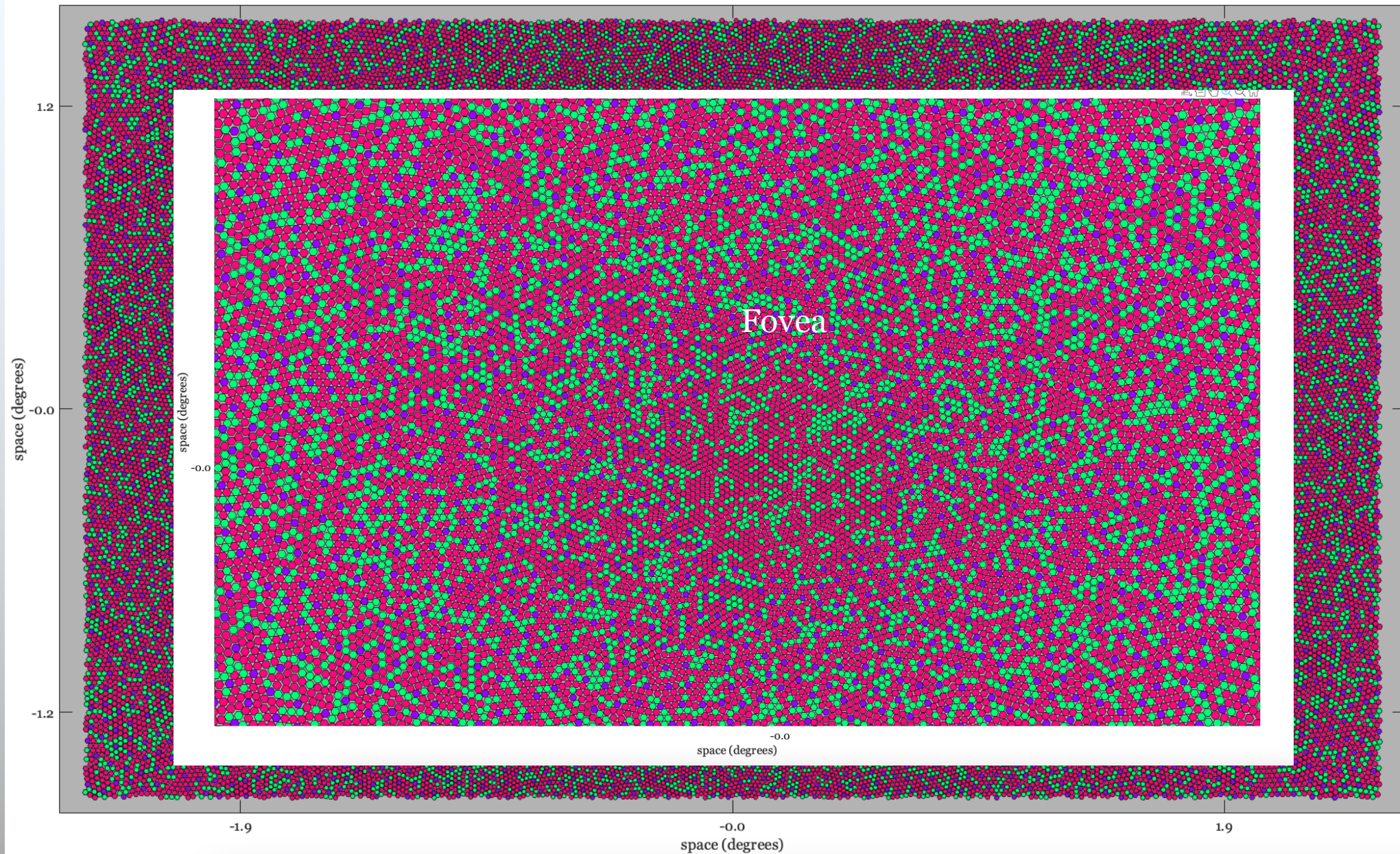
- The Navarro eye models enable us to model the change of accommodation (focus) for the human eye using complex scenes – or simple scenes!!
- And to calculate the excitations at the cone mosaic
- We will get to eye movements a little later



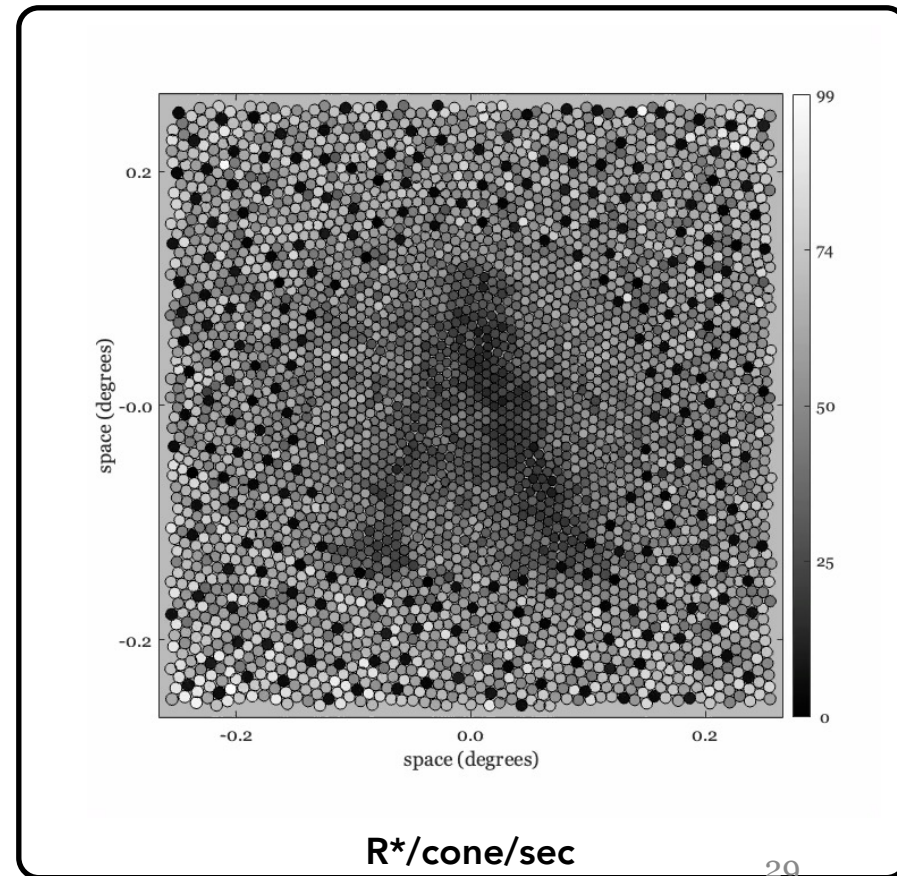
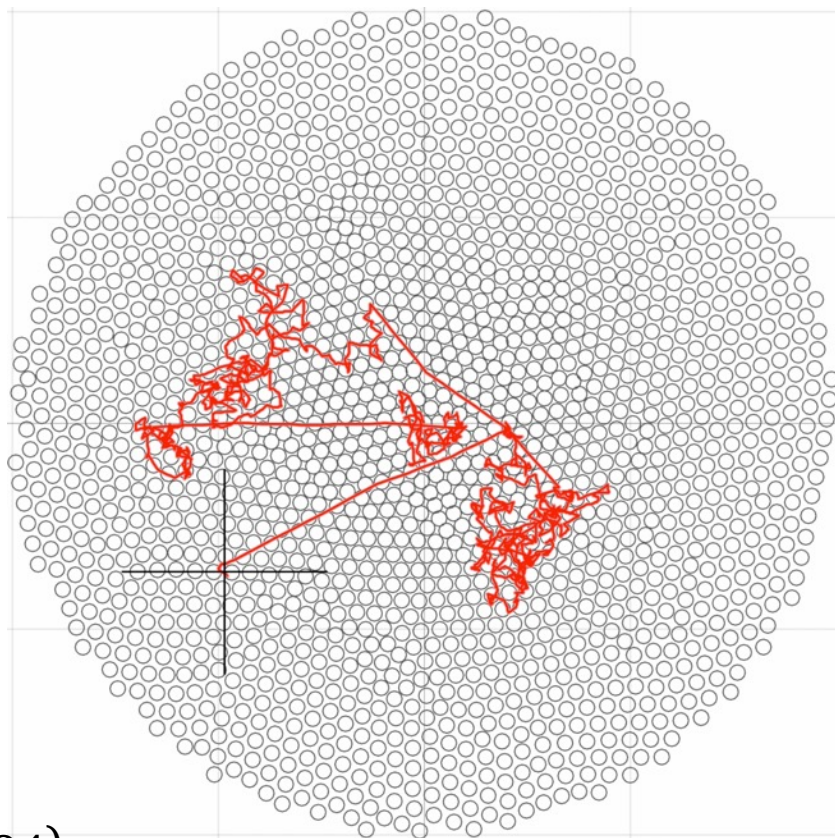
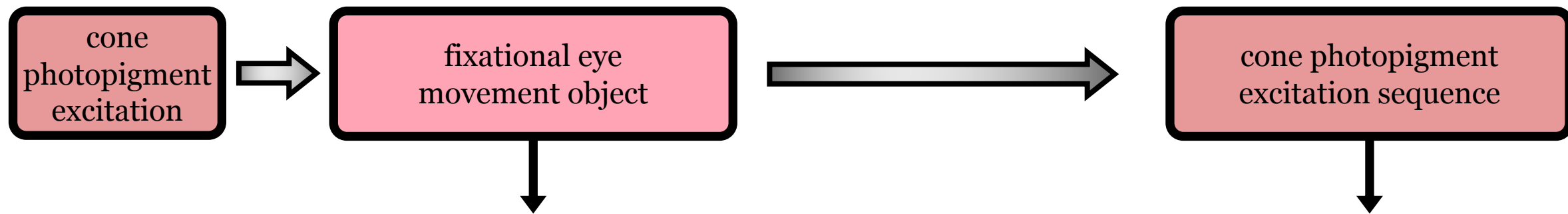
# The size and spatial arrangement of the cone mosaic

Retinal receptor sampling is shift  
VARYING

This slows down  
many  
computations

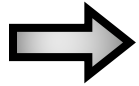


# Accounting for absolute sensitivity: fixational drift (Engbert and Kliegl, 2004)

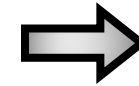


# Time: Photocurrent transduction (Angueyra & Rieke, 2013)

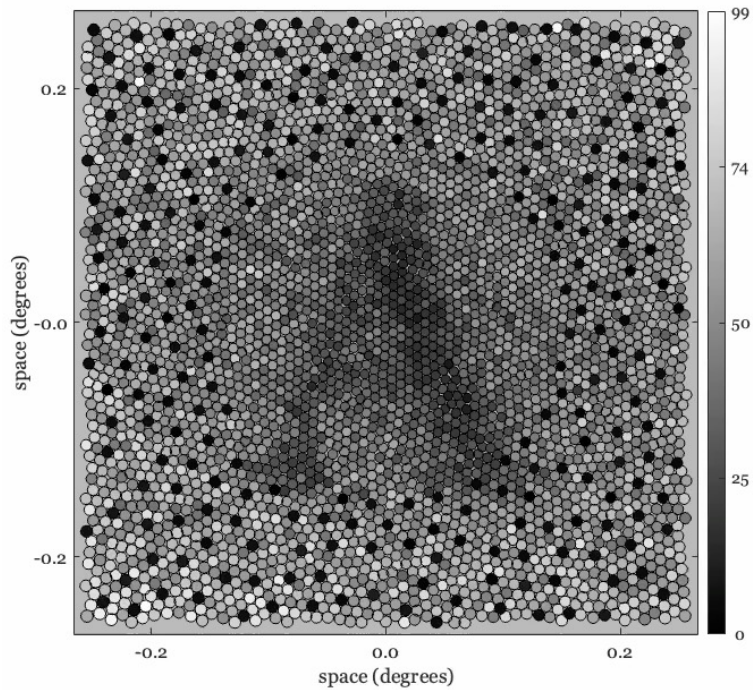
cone photopigment  
excitation sequence



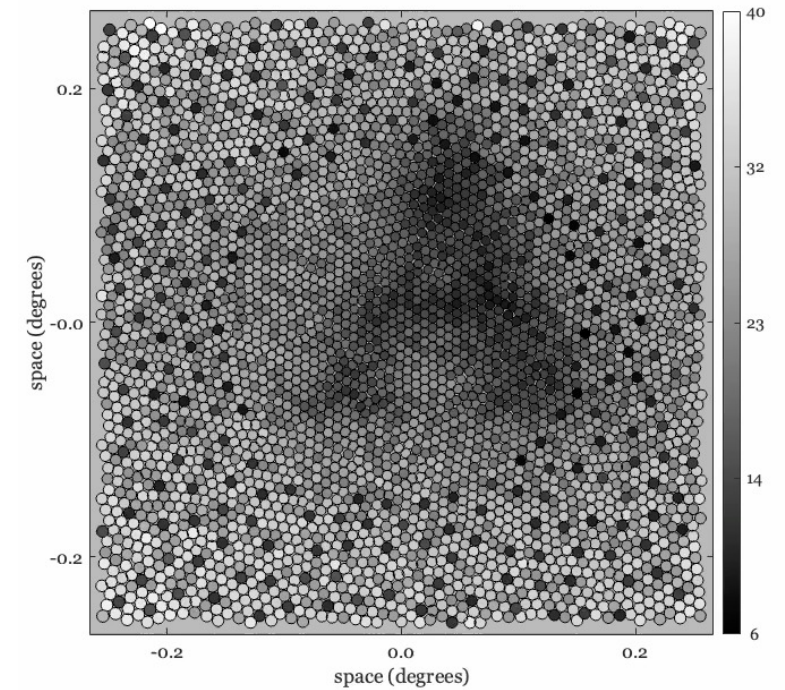
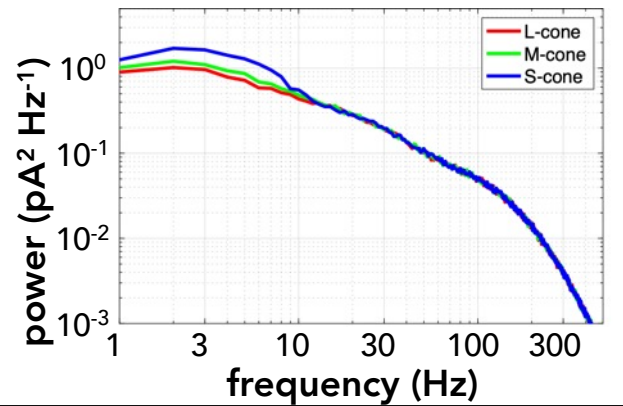
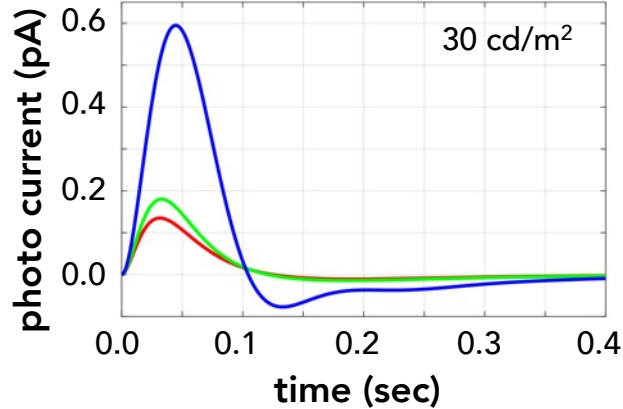
outer segment  
object



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 10 c/deg)
2. **Computational** modeling by observing neural responses shows a sensitivity drop across a range (accumulated factor: 2-3)
3. Inclusion of **fixation** and **linear computational** models shows sensitivity across a range (accumulated factor: 7-10).
4. Inclusion of **photocurrent encoding** further reduces sensitivity approaching psychophysical limits (accumulated factor: 18-30).



April 2019  
Volume 19, Issue 4  
ISSUE

OPEN ACCESS

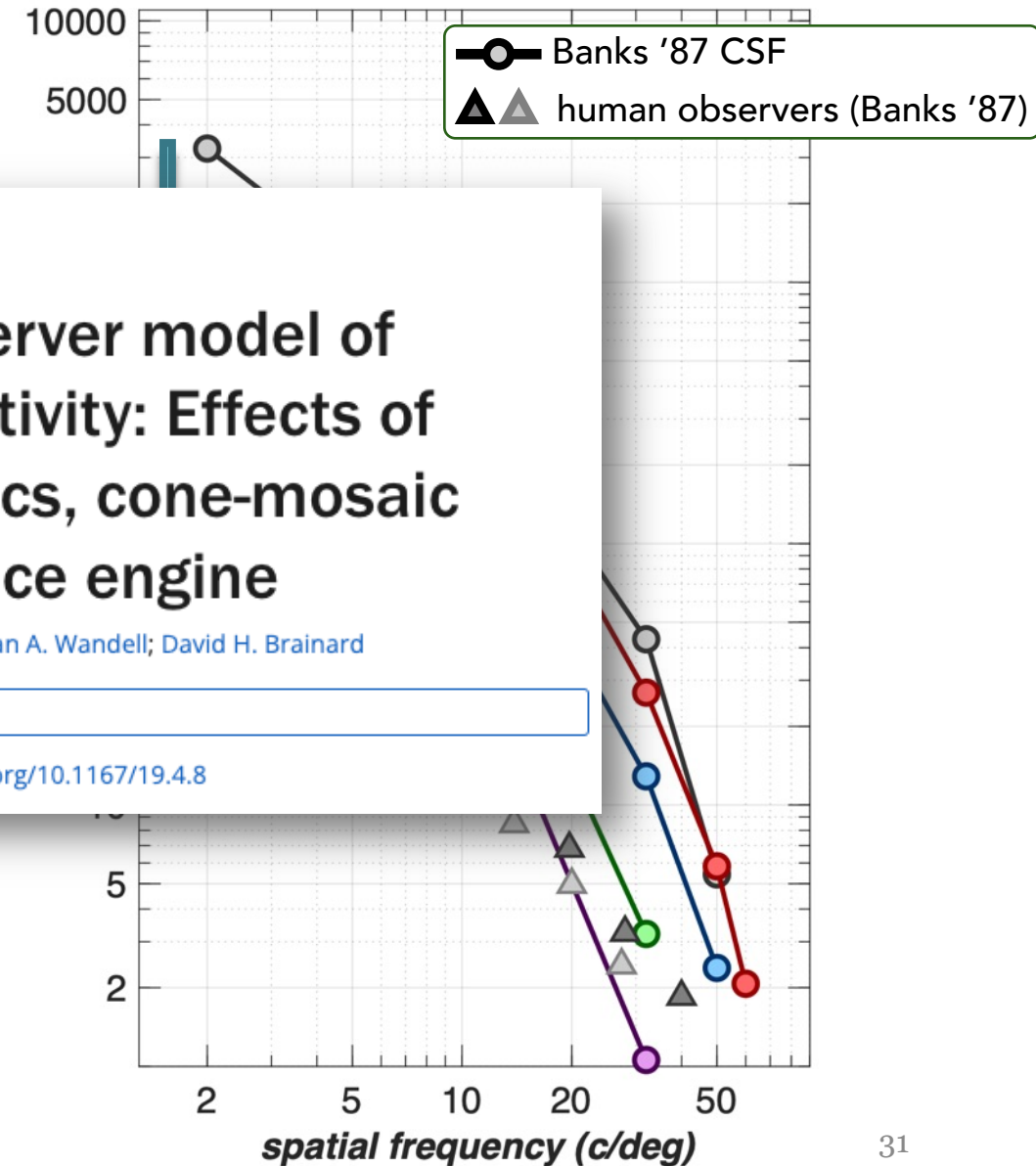
Article | April 2019

## A computational-observer model of spatial contrast sensitivity: Effects of wave-front-based optics, cone-mosaic structure, and inference engine

Nicolas P. Cottaris; Haomiao Jiang; Xiaomao Ding; Brian A. Wandell; David H. Brainard

+ Author Affiliations

Journal of Vision April 2019, Vol.19, 8. doi:<https://doi.org/10.1167/19.4.8>



# Biological data synthesis for a single performance curve (e.g., CSF)

- We must manage time series ( $\Delta t \sim 5$  ms) to account for eye movements and photocurrent
- The computations are NOT shift-invariant, making fft/convolutions inappropriate
- We must manage multiple experiments, each with 1K time series, to sweep out a performance curve

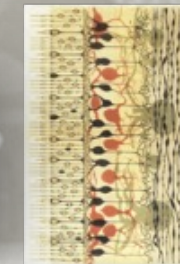
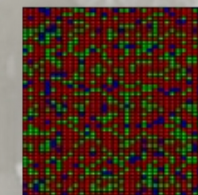
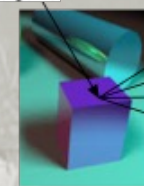
Category	Parameter	Value
Retinal irradiance		
	Spatial samples	36,000,000 (6 mm x 6 mm)
	Wavelengths	31 wavebands
	Time samples	200 (5 ms, 1 sec)
	Memory	200 GB
Rendering parameters		
	Objects	1000
	Polygons per asset	10,000 – 1,000,000
	Rays per pixel	2000
Cone images per MTF curve		
	Cone samples per image	2,000,000
	Time samples	200
	Number of time series	1000
	Memory	400 GB

## Third point

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

# Ongoing and future plans

- ISETCam/Bio: Ideas about data sharing for these scenes and the development of neural networks for automated driving assistance systems (ADAS)
- ISETCam: Increase performance for sensors with larger number of spatial samples, mixed pixel arrays, and complex color filter arrays
- ISETCam: Microlens computations and light fields for large sensor arrays
- ISETBio: Managing multiple neuronal systems (RGC, cortex);
- ISETBio: Subject variability; integration with image quality data



# 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

