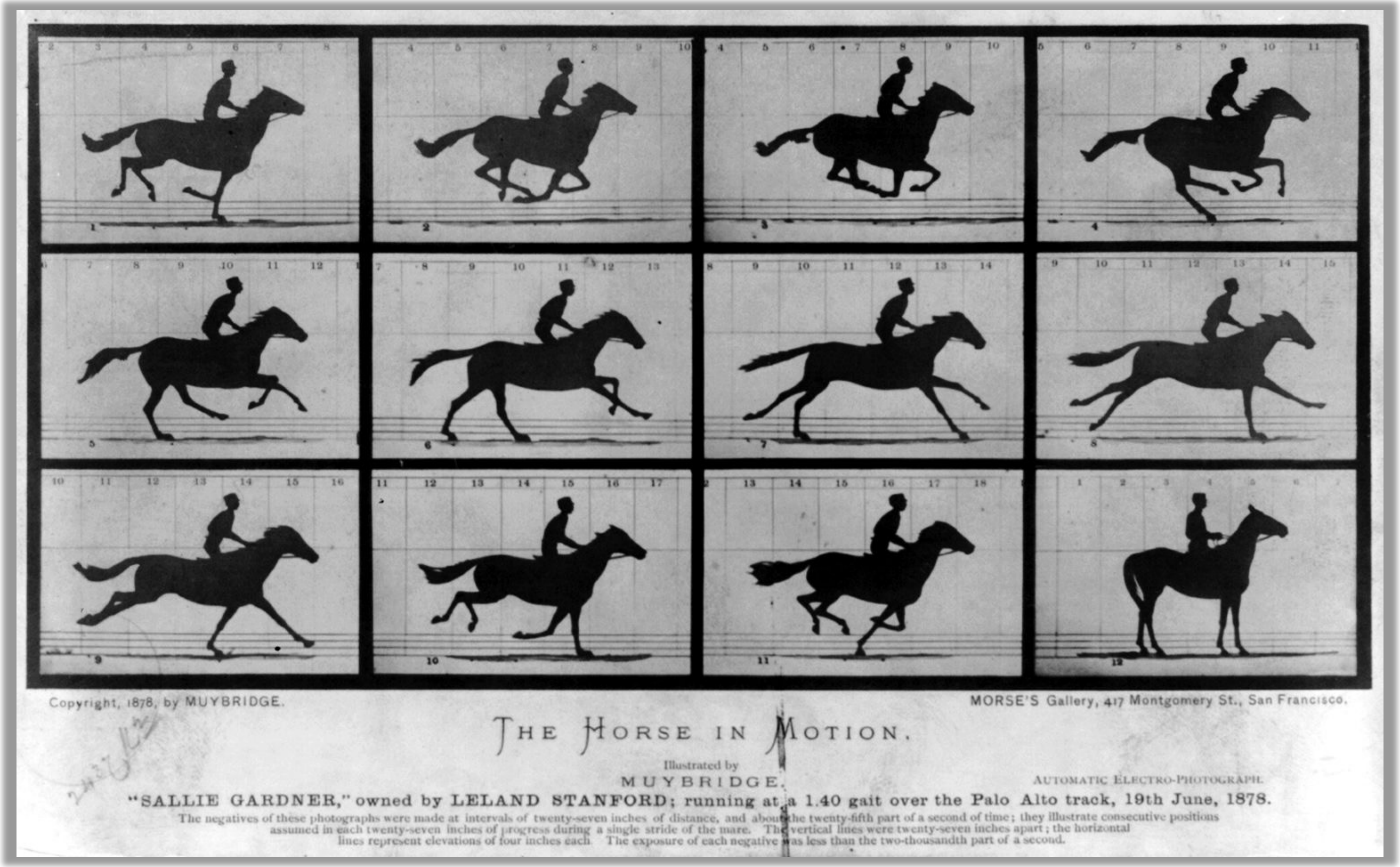
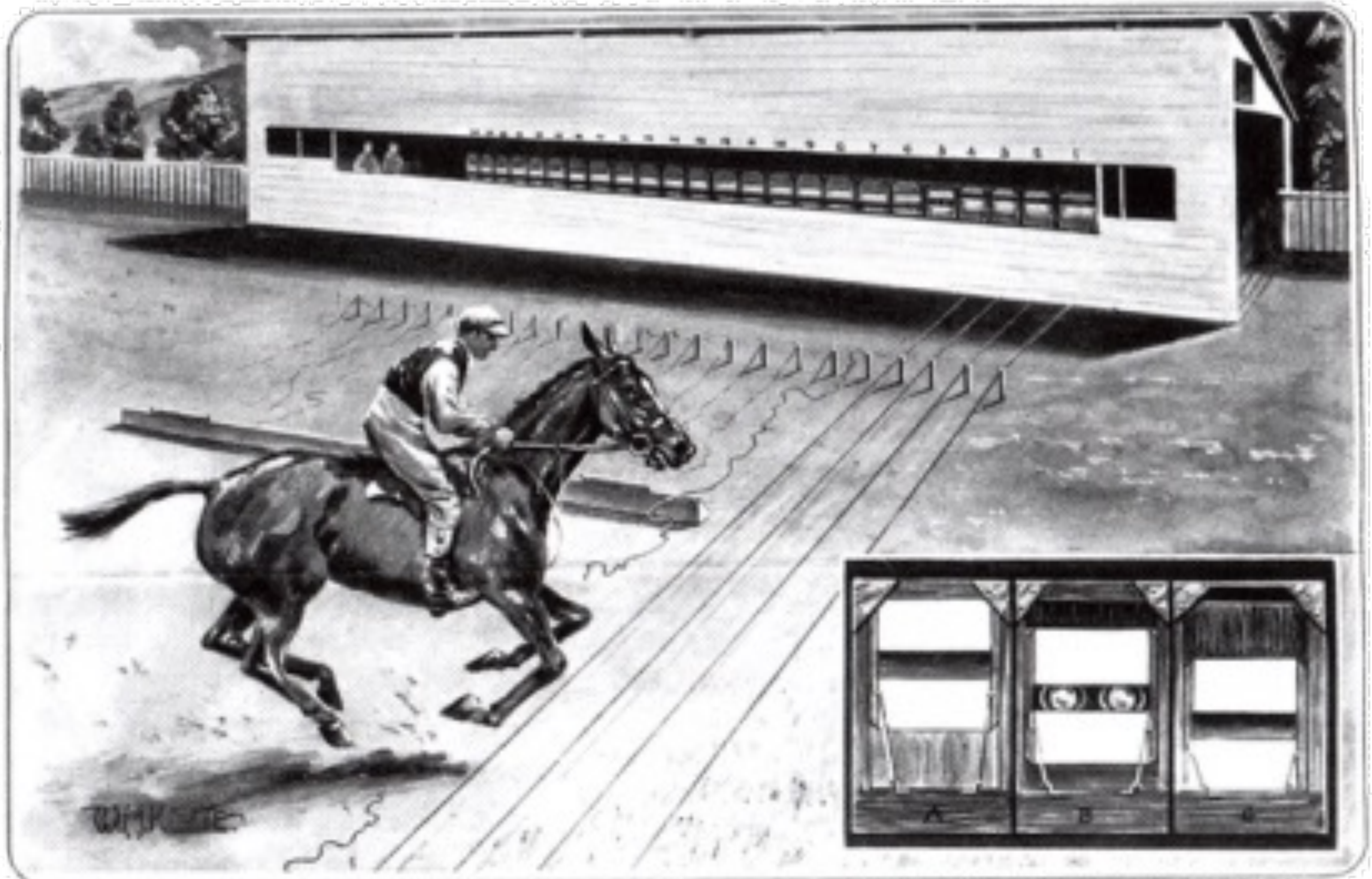
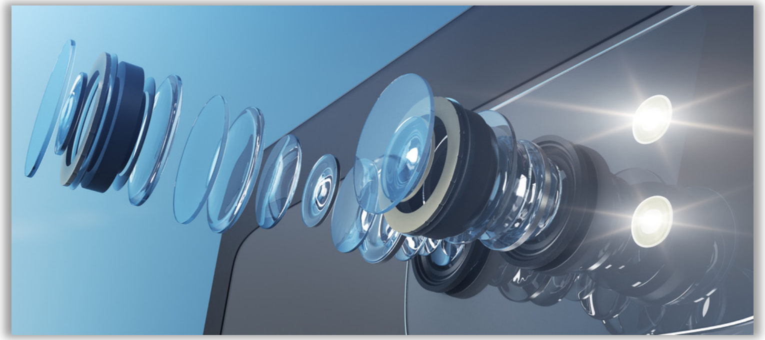
Course Introduction/Human Visual System

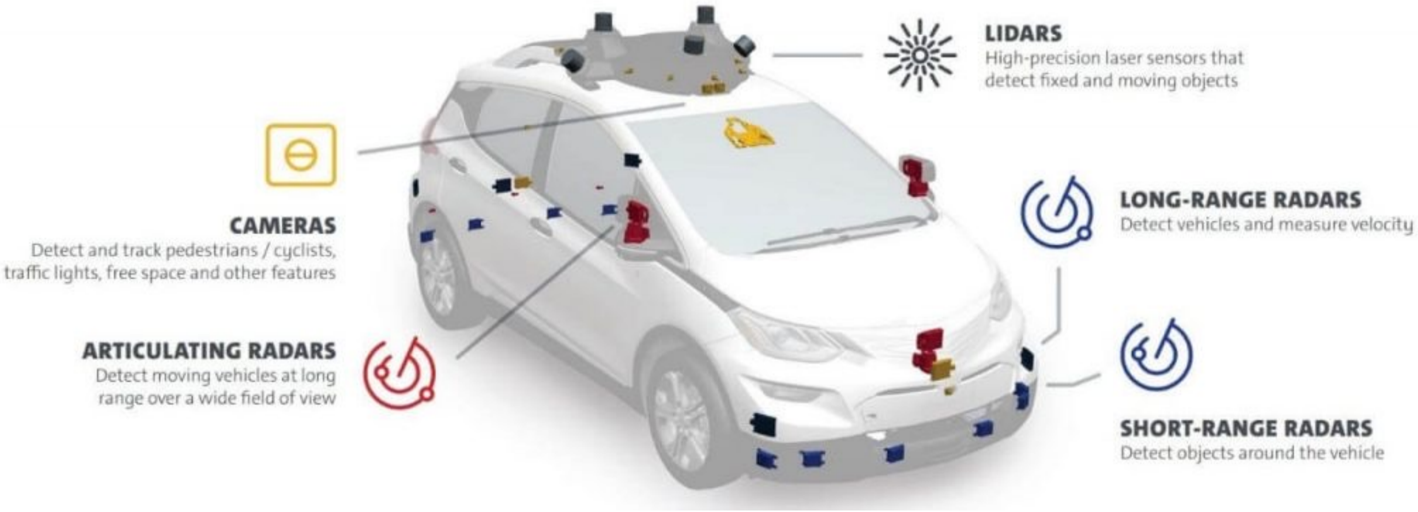
CSC2529
David Lindell
University of Toronto
cs.toronto.edu/~lindell/teaching/2529


Copyright, 1878, by MUYBRIDGE.
MORSE'S Gallery, 417 Montgomery St., San Francisco.
THE HORSE IN MOTION.
Illustrated by
MUYBRIDGE.
AUTOMATIC ELECTRO-PHOTOGRAPH.
"SALLIE GARDNER," owned by LELAND STANFORD; running at a 1.40 gait over the Palo Alto track, 19th June, 1878.
The negatives of these photographs were made at intervals of twenty-seven inches of distance, and about the twenty-fifth part of a second of time; they illustrate consecutive positions assumed in each twenty-seven inches of progress during a single stride of the mare. The vertical lines were twenty-seven inches apart; the horizontal lines represent elevations of four inches each. The exposure of each negative was less than the two-thousandth part of a second.

Muybridge's Multi-Camera Array






LIDARS
High-precision laser sensors that detect fixed and moving objects
CAMERAS
Detect and track pedestrians / cyclists, traffic lights, free space and other features
ARTICULATING RADARS
Detect moving vehicles at long range over a wide field of view
LONG-RANGE RADARS
Detect vehicles and measure velocity
SHORT-RANGE RADARS
Detect objects around the vehicle


Nikon

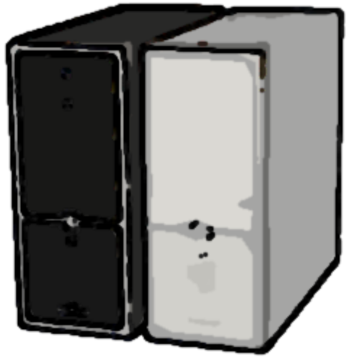

SIGMA

Image: National Geographic

What is Computational Imaging?

+

+



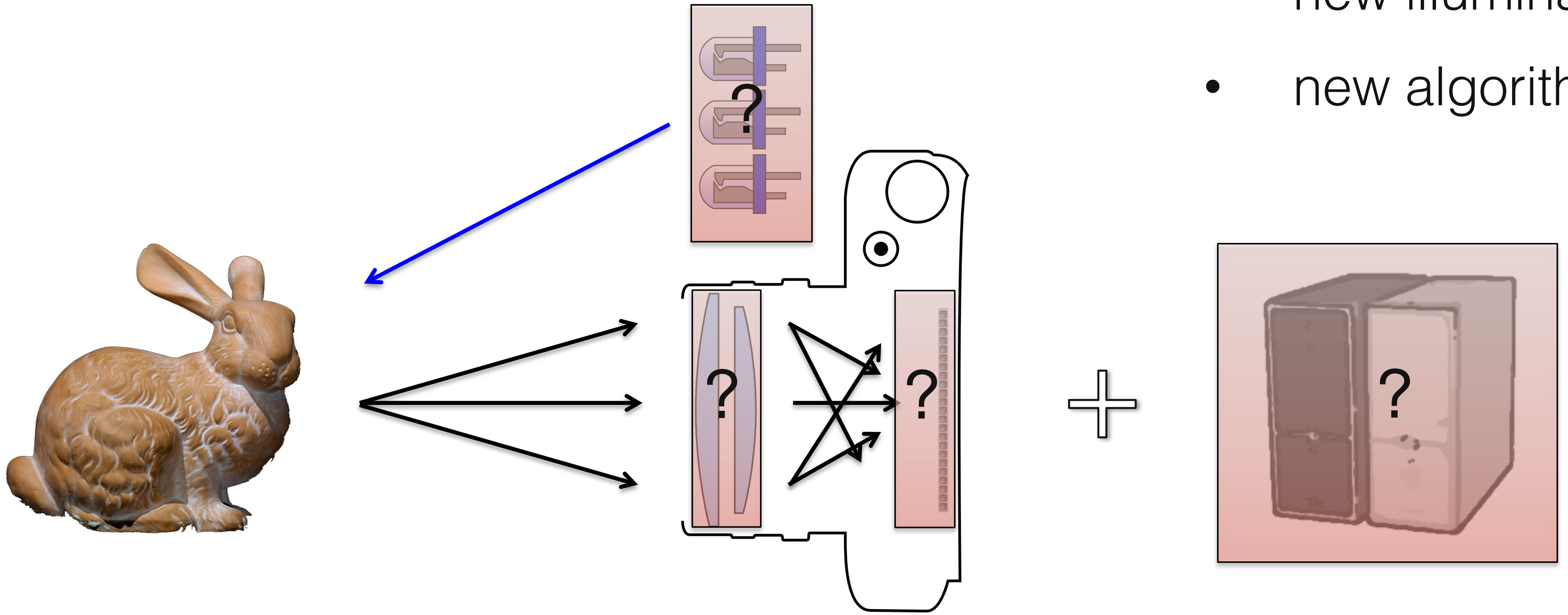
optics sensing computation

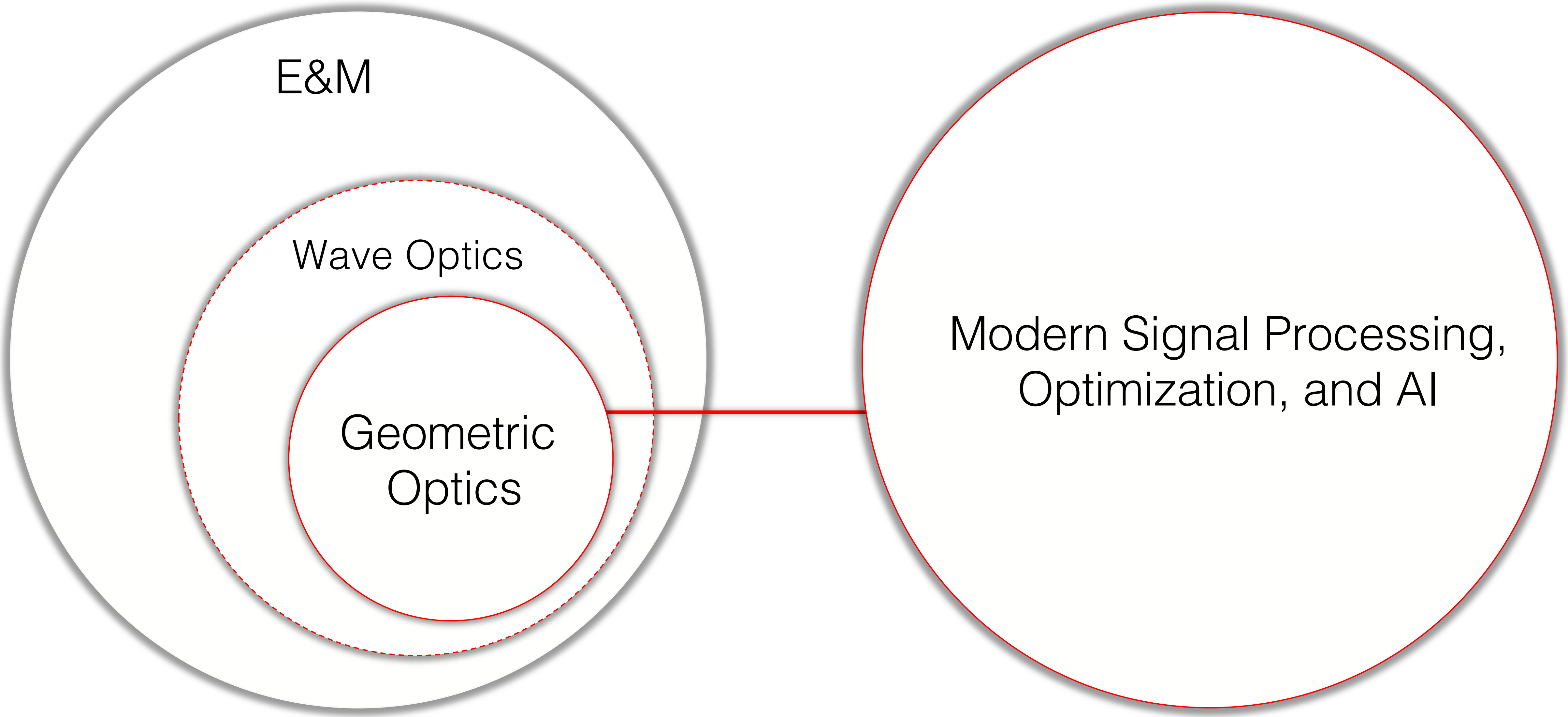
Computational Imaging

What is Computational Imaging?

1. optically encode scene information
2. computationally recover information

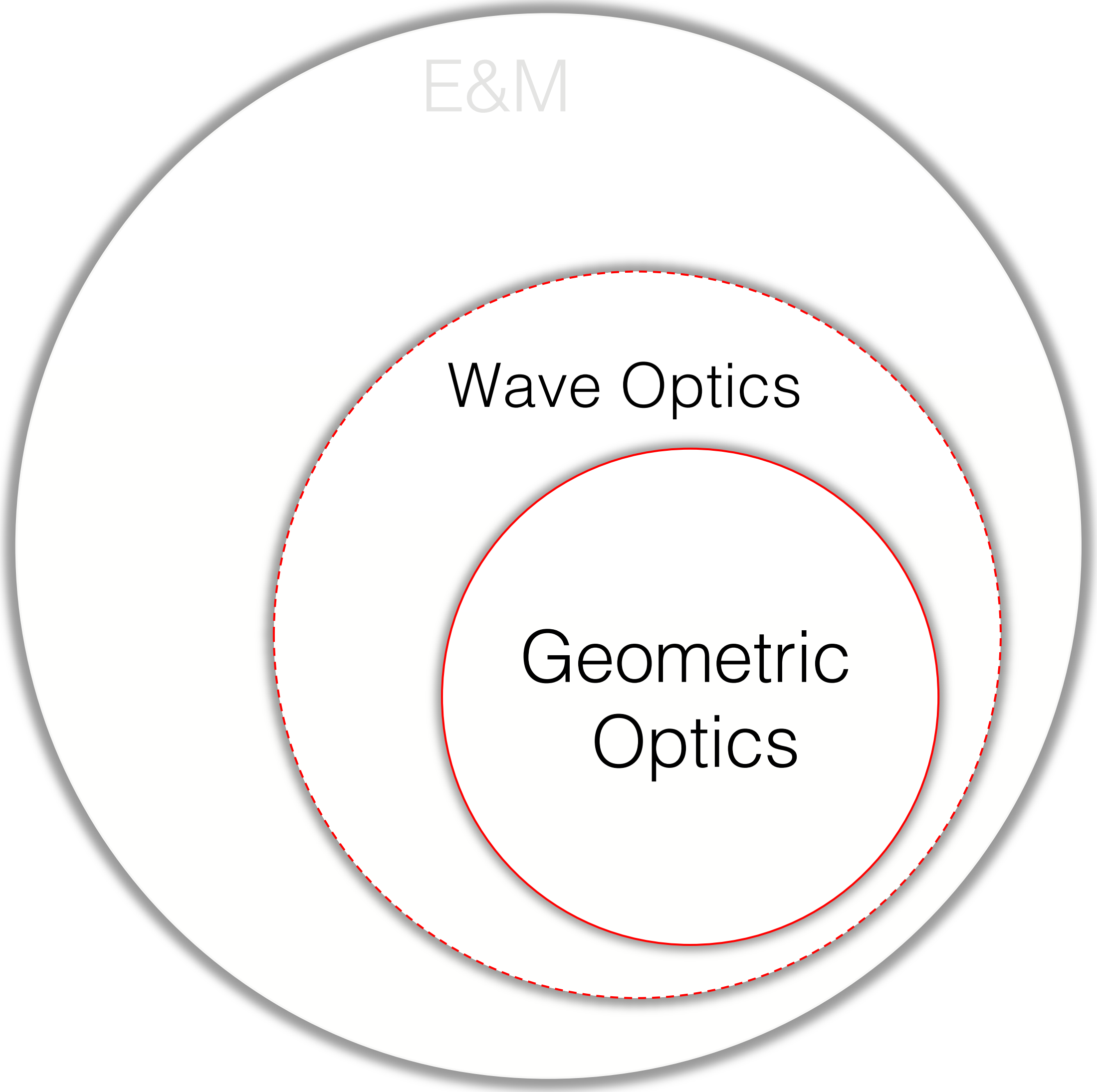
- new optics
- new sensors
- new illumination
- new algorithms





E&M
Wave Optics
Geometric Optics
Modern Signal Processing, Optimization, and AI

What is Light?



- light as rays
- unit: (spectral) radiance
- properties: wavelength, polarization, direction, …
- only brief introduction & outlook for wave optics

Course Fast Forward

Recording Notice

- Lectures and Problem Sessions in this course are recorded and published to Quercus

- If you ask a question your voice may be recorded

Acknowledgments

- Lecture material adapted from EE367: Computational Imaging by Gordon Wetzstein at Stanford University
- Materials also build on work by many others: Marc Levoy, Fredo Durand, Ramesh Raskar, Shree Nayar, Paul Debevec, Kyros Kutulakos, Matthew O'Toole

Instructors

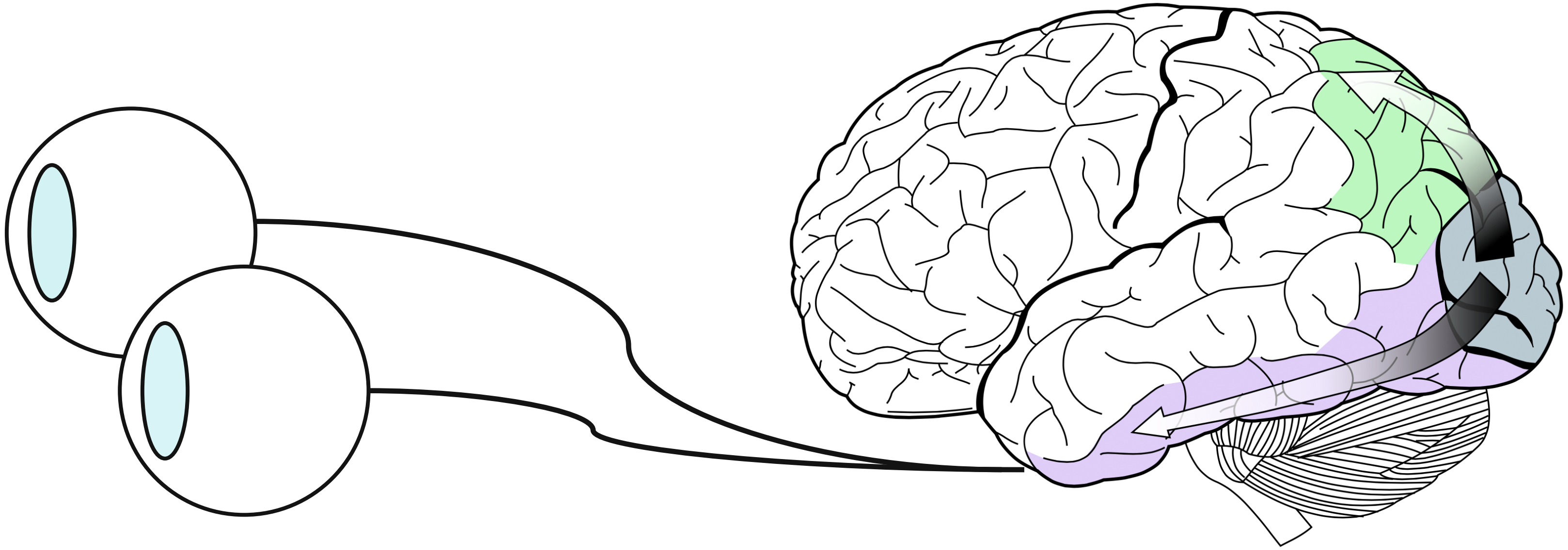
David Lindell

Parsa Mirdehghan

Anagh Malik

The Human Visual System

- anatomy of the eye
- acuity, color, 3D vision
- contrast sensitivity
- conflicts in displays
- refractive errors



wikipedia

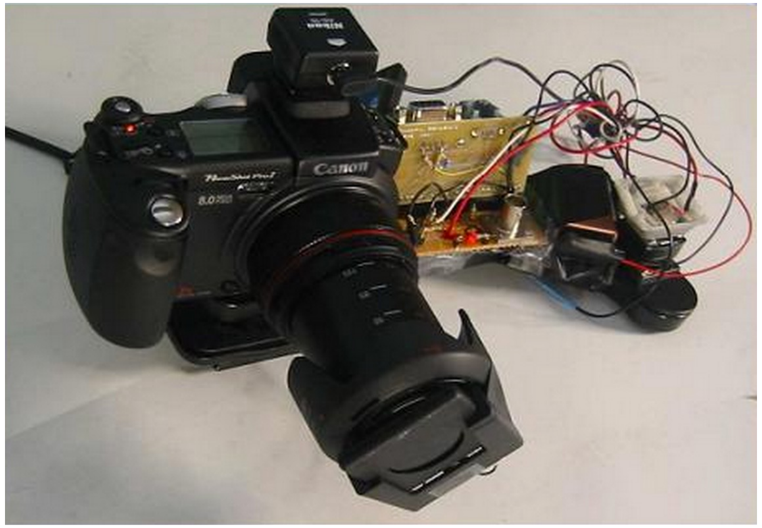
sensors ⟶ network ⟶ compute

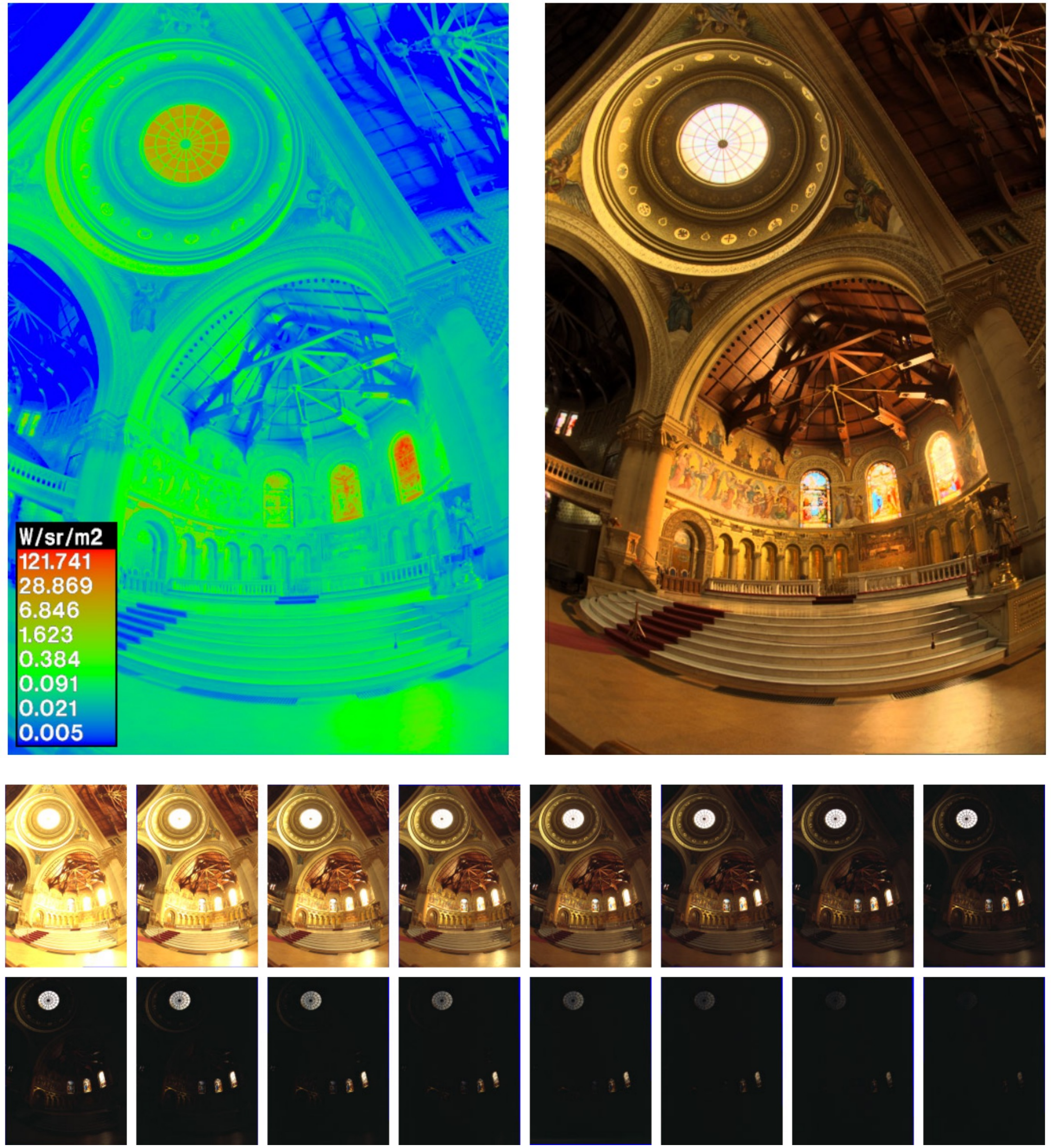
Digital Photography

- optics
- aperture
- depth of field
- field of view
- exposure
- noise
- color filter arrays
- imaging processing pipeline

Computational Photography

- High-dynamic range imaging
- Tone mapping
- Burst photography
- Coded apertures
- …

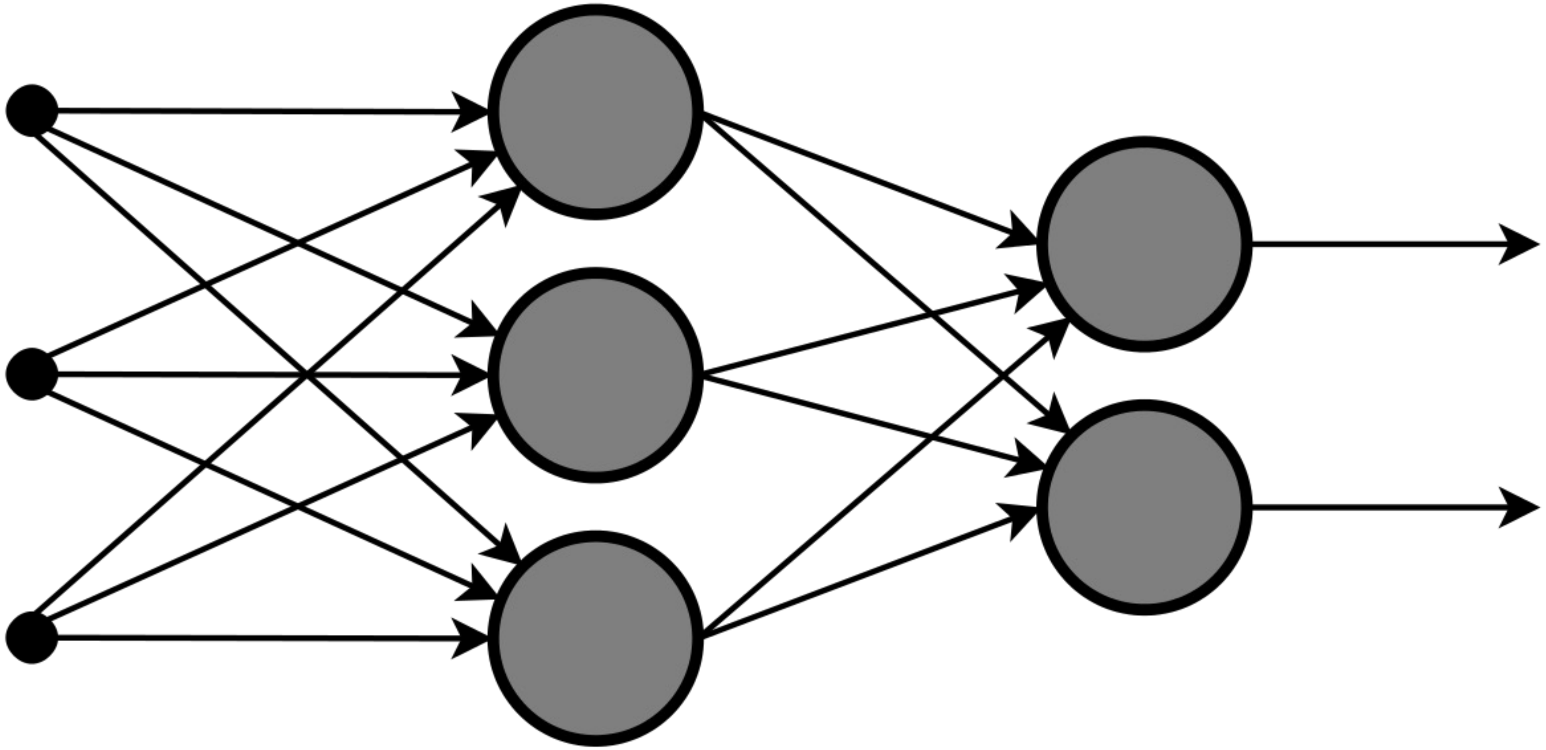




Debevec & Malik, 1997

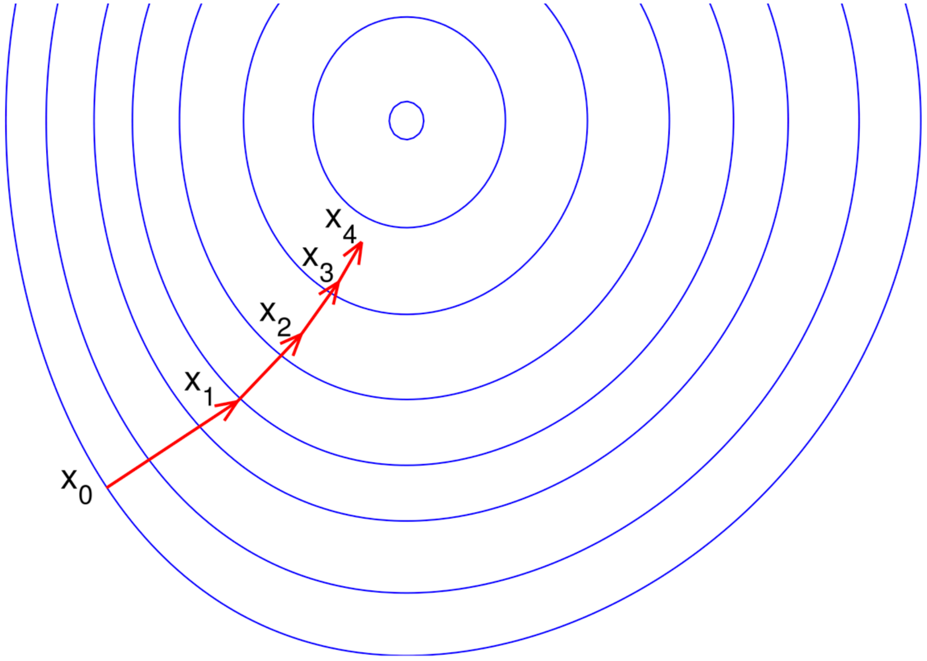
Deep Learning for Computational Imaging

- Convolutional neural networks
- DnCNN
- U-Net
- …



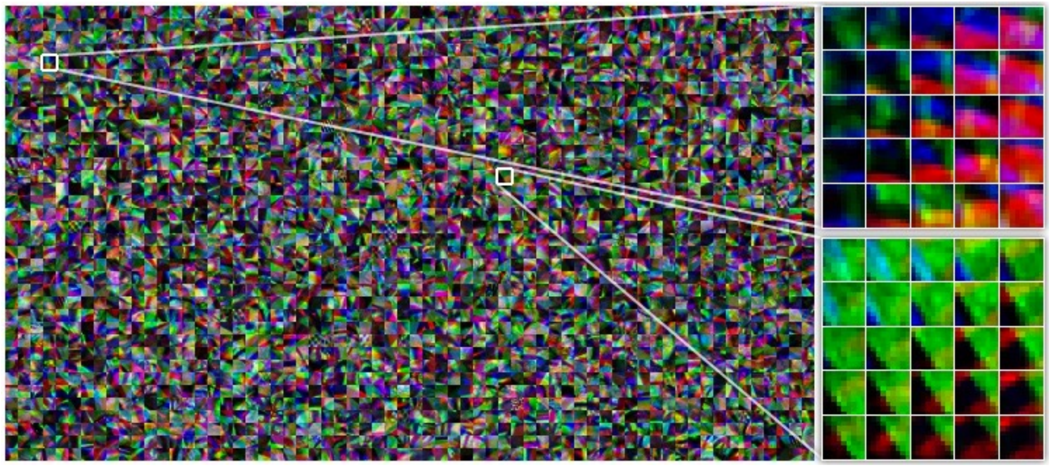
Optimization & Deep Learning

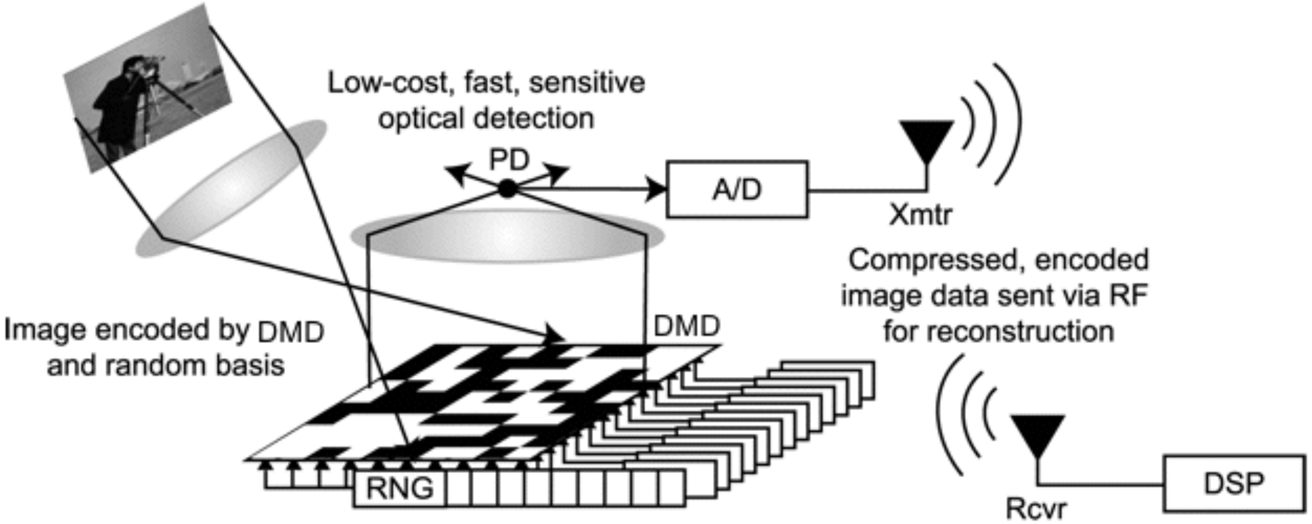
- Non-linear optimization
- Proximal gradient methods (ADMM)
- Iterative optimization with deep priors
- Solving general inverse problems in imaging
- …



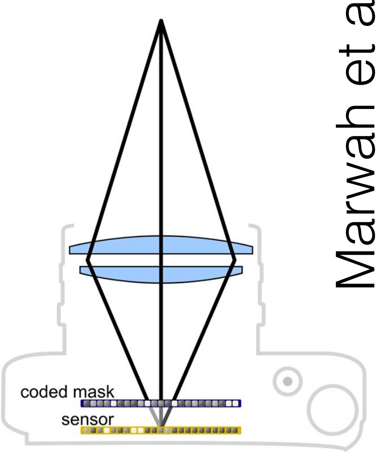
Compressive Imaging

- single pixel camera
- compressive hyperspectral imaging
- compressive light field imaging
- …





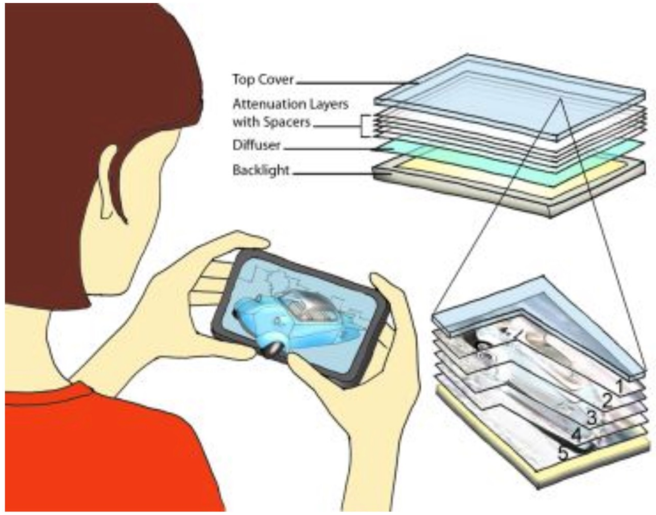
Wakin et al. 2006



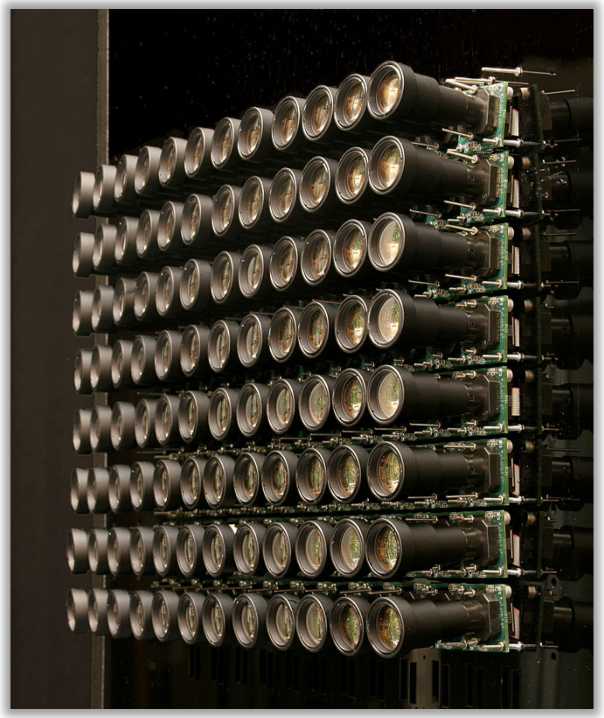
Marwah et al., 2013

Light Field Imaging

- Plenoptic function
- Light fields
- 3D displays
- …



[Wetzstein et al. 2011]



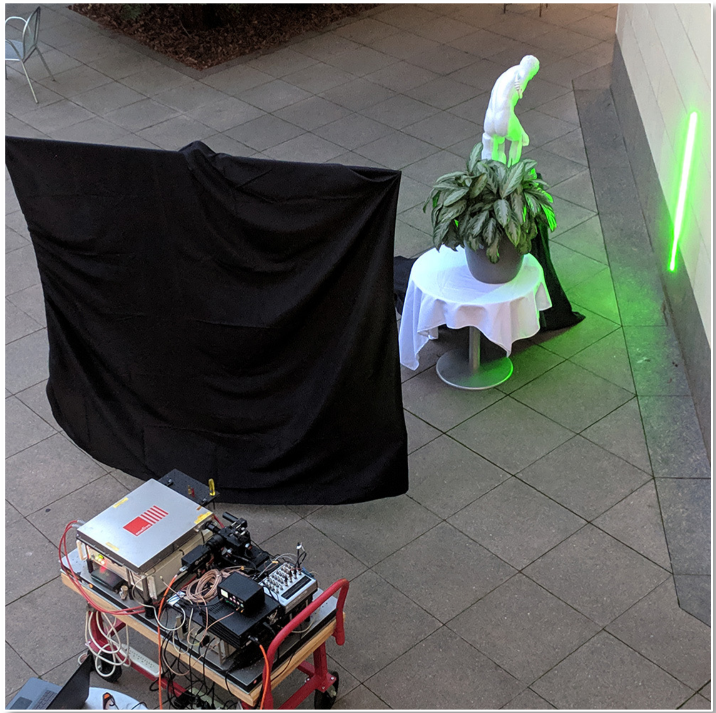
[Wilburn et al. 2005]

Lytro Illium

Time-of-Flight Imaging

- Lidar
- Single-photon imaging
- Non-line-of-sight imaging
- …

Velodyne

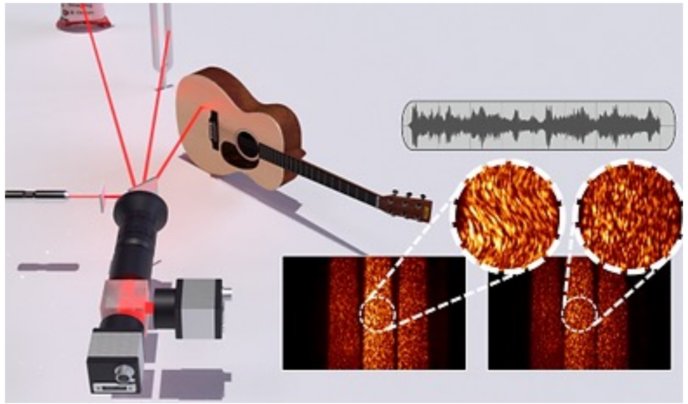


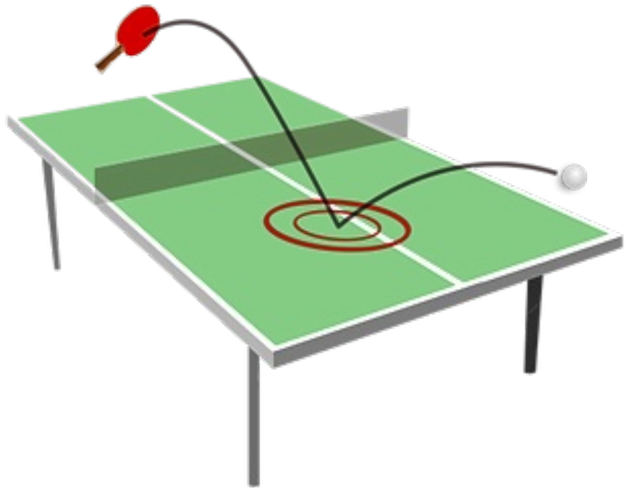
[Lindell et al. 2019]

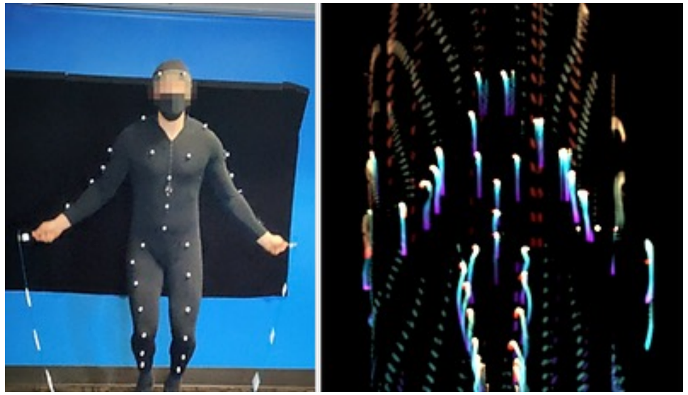
[O'Toole et al. 2017]

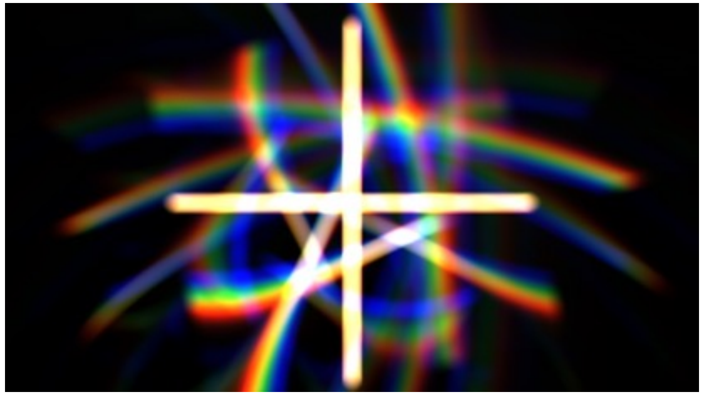
Guest Lecture: Mark Sheinin
Computational Imaging for Enabling Vision Beyond Human Perception

Mark Sheinin
Incoming Asst. Professor
Weizmann Institute



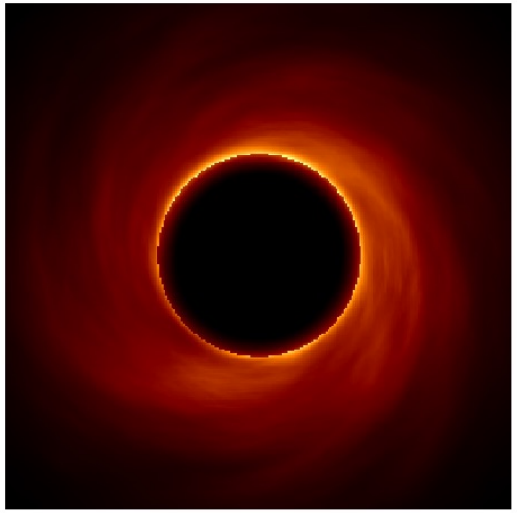


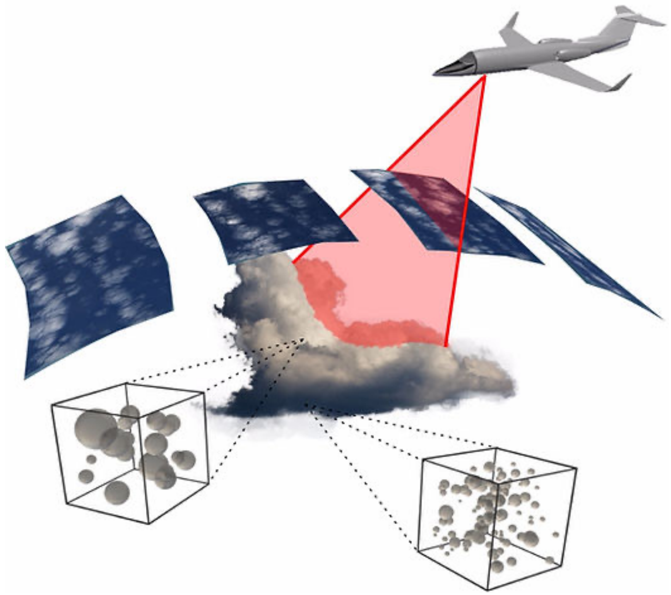


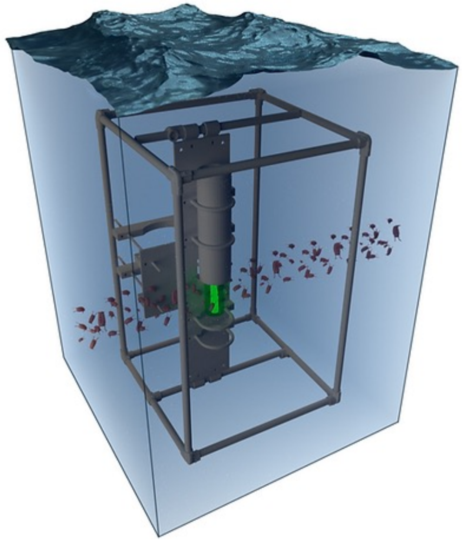


Guest Lecture: Aviad Levis

Aviad Levis
Incoming Asst. Professor
University of Toronto







Class Details

(no formal) Prerequisites (but …)

- strong *programming skills*, ideally Python
- *linear algebra*
- basic knowledge of *Fourier transforms*
- maybe a bit of (statistical) signal processing, but not absolutely required
- basic computer graphics or computer vision could be helpful, but also not required

Related, Possibly Helpful Classes

UofT Classes:

CSC2530	Computational Imaging and 3D Sensing
CSC2305	Numerical Methods for Optimization Problems
CSC2503	Foundations of Computer Vision
CSC2516	Neural Networks and Deep Learning
ECE537	Random Processes
ECE1505	Convex Optimization
ECE1512	Digital Image Processing and Applications

Requirements and Grading

- 6 assignments: 50%
- major final project (teams of ≤ 3): 50%
  - discuss project ideas with TA & instructor!
  - project proposal due: 15/11, 11:59pm
  - final presentation: 7/12, 2-4pm
  - reports and source code due: 7/12, 11:59pm

Resources (see course website!)

- website: cs.toronto.edu/~lindell/teaching/2529/
- contact: csc2529-fall2324-staff-l@listserv.utoronto.ca
- office hours (TA, problem sessions): Tues 1:00–2:00pm, BA5256
- office hours (Instructor, projects): Mon 4:00–5:00pm BA7228
- Ed Discussion (see Quercus for link)

Tentative Schedule

cs.toronto.edu/~lindell/teaching/2529/

What we don’t discuss

- no medical imaging, but same concept apply – medical imaging projects are encouraged!
- outlook on wave optics / diffractive imaging but not focus on this topic

Lectures and Problem sessions

- 1 lecture per week: Mon. 2-4pm in Galbraith 120 in person (recording will be available on Quercus after class)

- 1 problem session (first 6 weeks): Tues BA5256 (recording will be available on Quercus after class)

- attendance strongly recommended, but everything is recorded

Assignments

- 6 assignments: mix of theory, programming, and HW1 has a bit of hands-on building
- out every Mon (starting this week), due Wed week after at 11:59pm (midnight)
- 3 late days for the term. Please don't ask for extensions unless something exceptional comes up.
- discussion among students encouraged, **but must submit own solution and acknowledge others that you discussed this with (no copying solutions)**
- submission via Quercus

Course Projects & Proposal

- individual or teams of up to 3 people
- 50% of your grade – plan on ~50-60 h per person!

- Nov 15: short project proposal = 1-2 pages with
  - motivation
  - related work
  - project overview
  - milestones, timeline & goals
  - at least 3 scientific references
  - we may ask you to revise the proposal, will assign a mentor to your team

Course Projects

- Thursday Dec 7: in-person project poster + demo session
  - see poster template on website
  - More details later

Course Projects

- Dec 7: report + source code due (at midnight)

- report = conference paper format ~6 pages with
  - abstract
  - introduction
  - related work
  - theory
  - analysis
  - results
  - discussion and conclusion
  - references
  - see latex template on website

Course Projects

- must also submit source code along with report!

- proposals, reports, source will be available on course website
  - only use non-copyrighted material
  - no projects that require NDA or company secrets
  - may request that source code / report may not be public – contact staff

Possible Course Projects

- be experimental!

- Image enhancement for under-display cameras
- Optimization or deep learning for your favorite inverse problem in imaging
- …

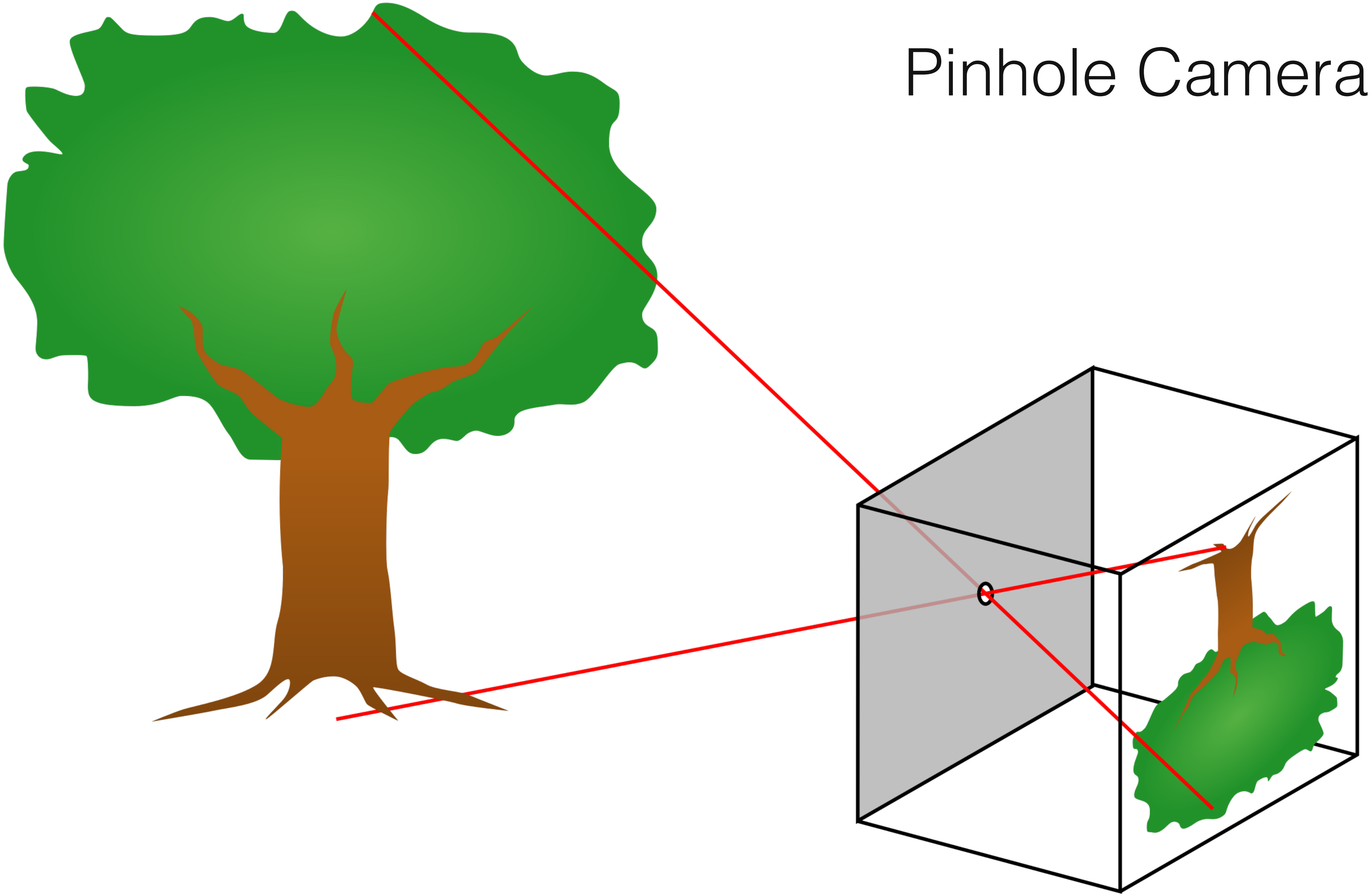
Possible Course Projects

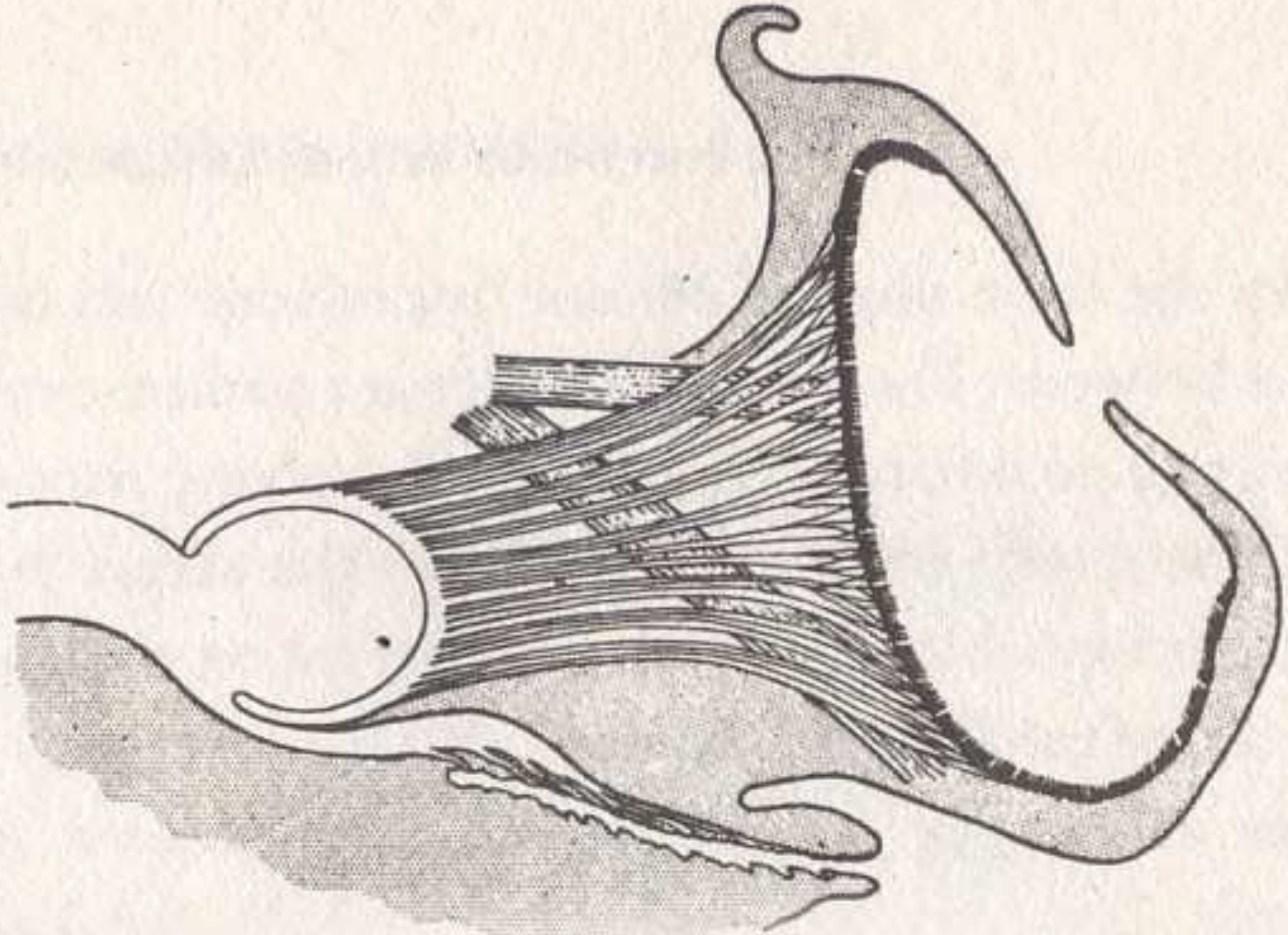
See previous course projects (proposals, reports, code, posters) on the course website!

The Human Visual System

nautilus eye, wikipedia

Pinhole Camera





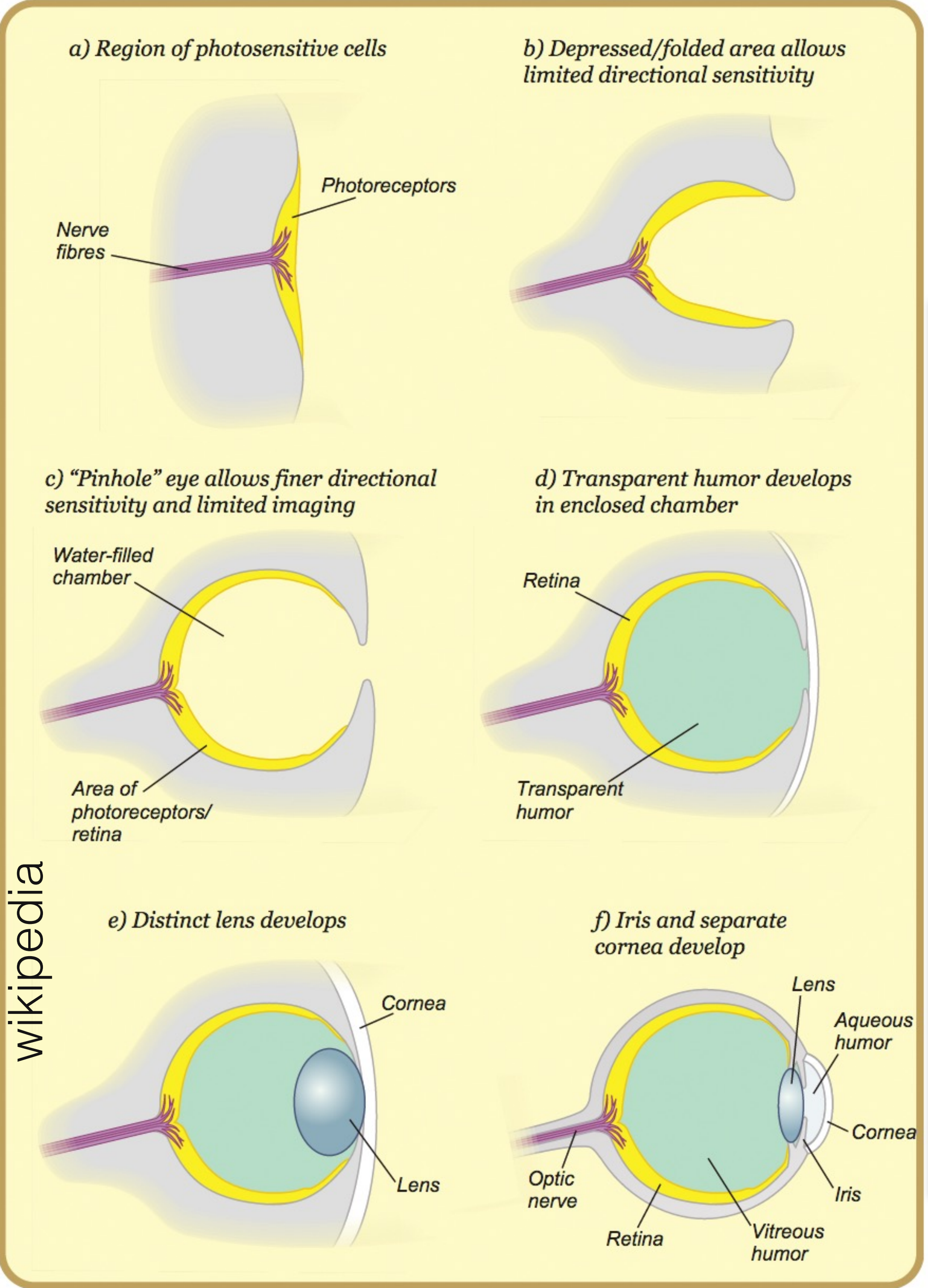
a

Figure 5.8 (opposite) A range of invertebrate eyes that illustrate approaches to the formation of crude but effective images: (a) *Nautilus*'s pinhole eye; (b) marine snail; (c) bivalve mollusc; (d) abalone; (e) ragworm.

Dawkins, "Climbing Mount Improbable", Norton & Company, 1997

reptile eye, http://pichost.me/1608580/

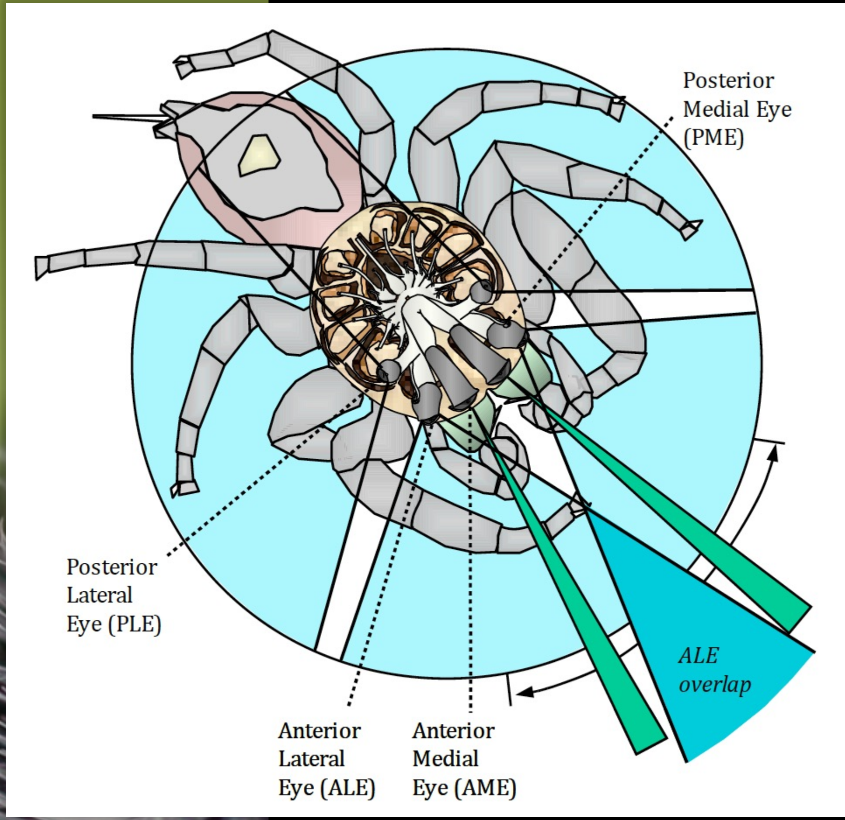
Evolution of the Eye



wikipedia

owl, https://www.pinterest.com/pin/452400725039917330/

pigeon, http://globe-views.com/dreams/pigeon.html



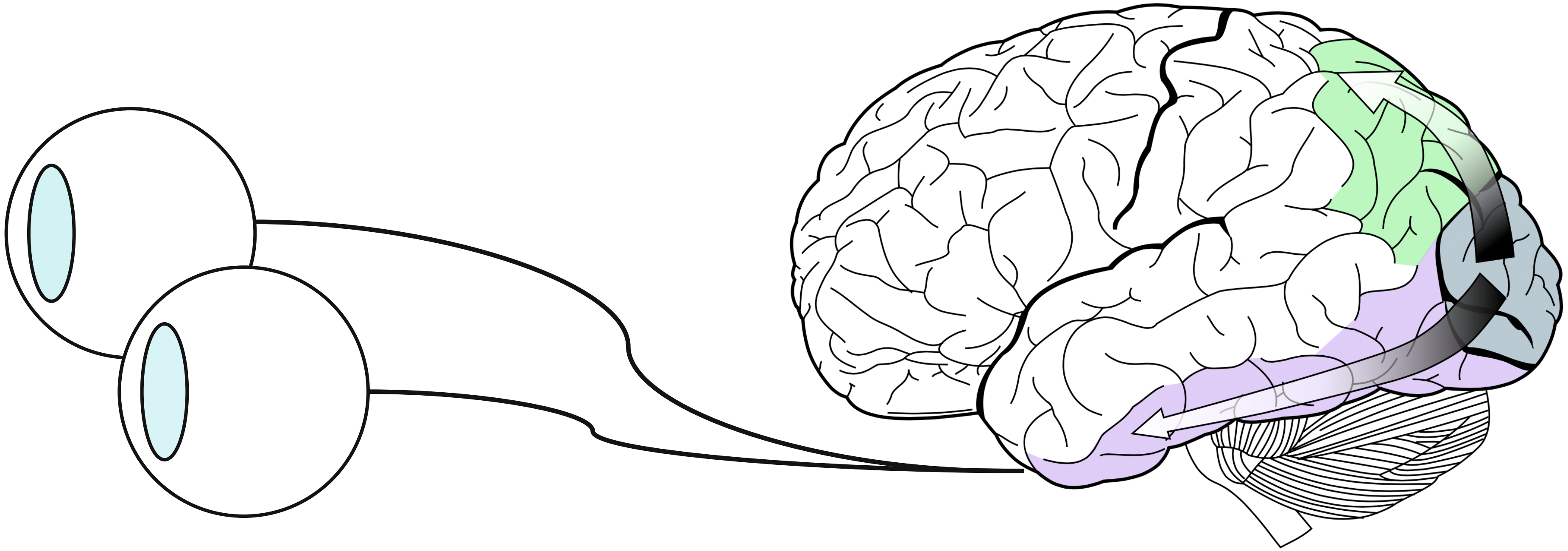
jumping spider, wikipedia

national geographics

Summary of Human Visual System (HVS)

- visual acuity: 20/20 is ~1 arc min
- field of view: ~190° monocular, ~120° binocular, ~135° vertical
- temporal resolution: ~60 Hz (depends on contrast, luminance)
- dynamic range: instantaneous 6.5 f-stops, adapt to 46.5 f-stops
- color: everything in the CIE xy diagram; distances are linear in CIE Lab
- depth cues in 3D displays: vergence, focus, conflicts, (dis)comfort
- accommodation range: ~8cm to ∞, degrades with age

Overview



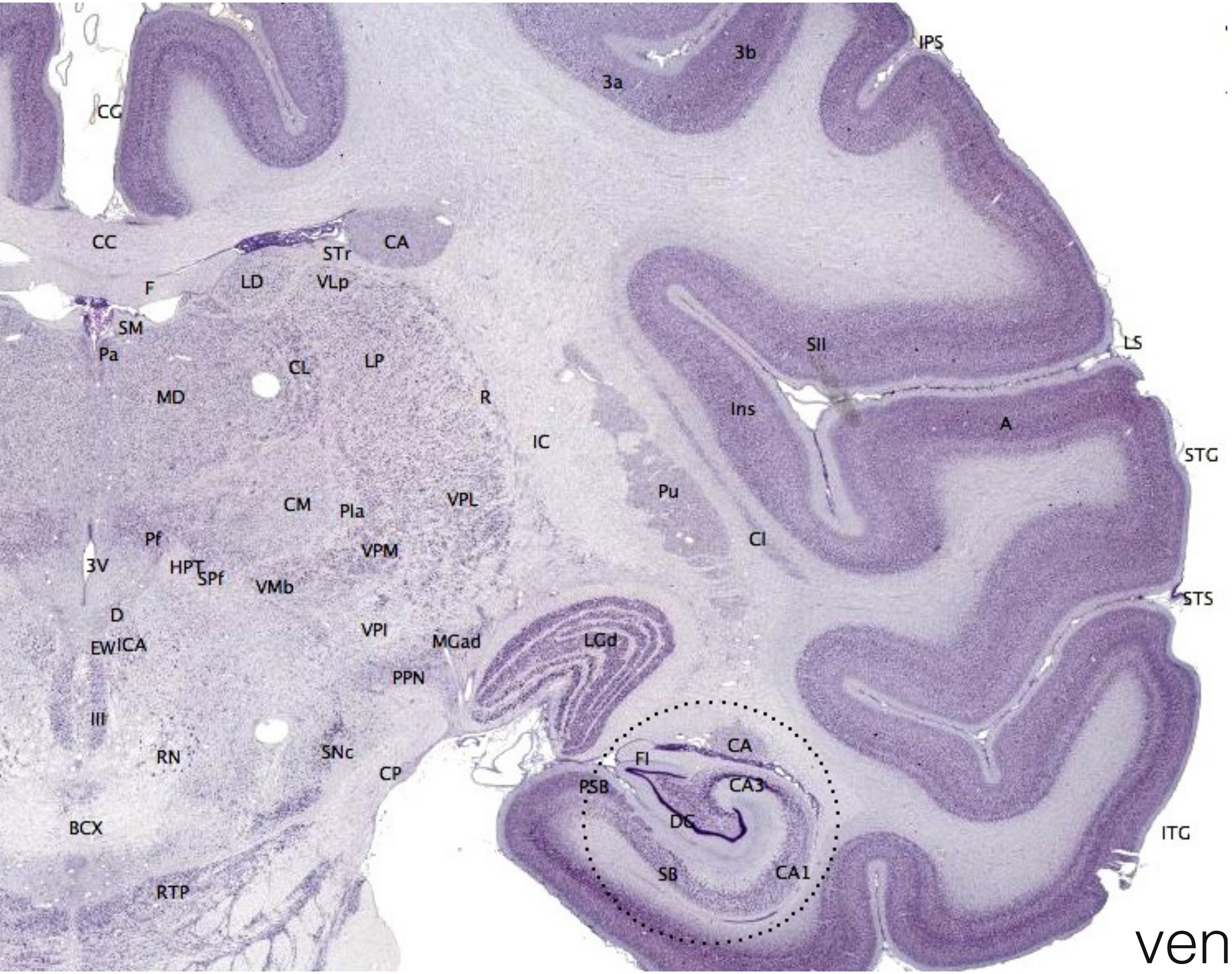
wikipedia

sensors → network → compute

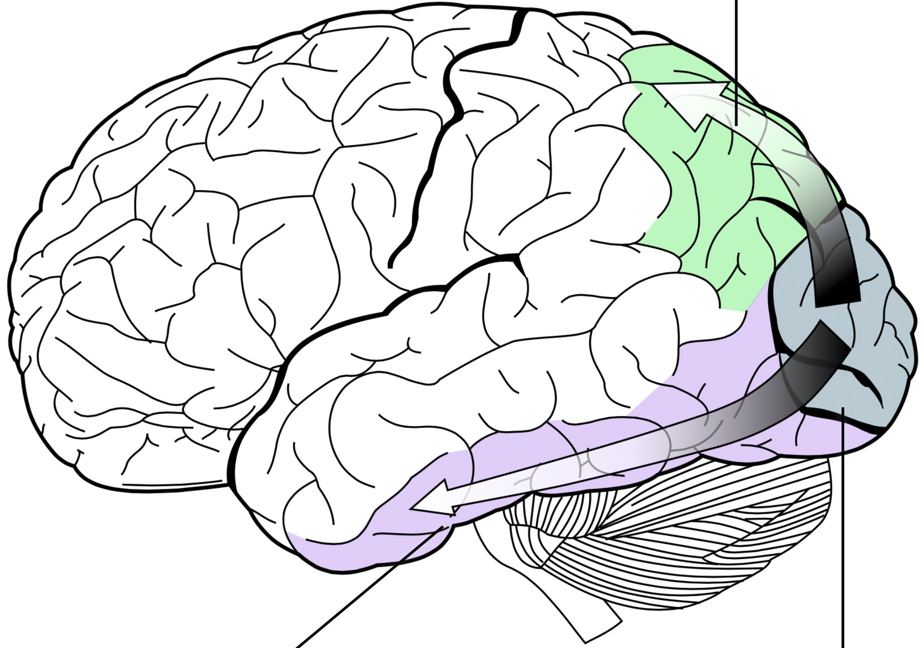
low-level processing

high-level processing

Overview



dorsal stream: spatial awareness

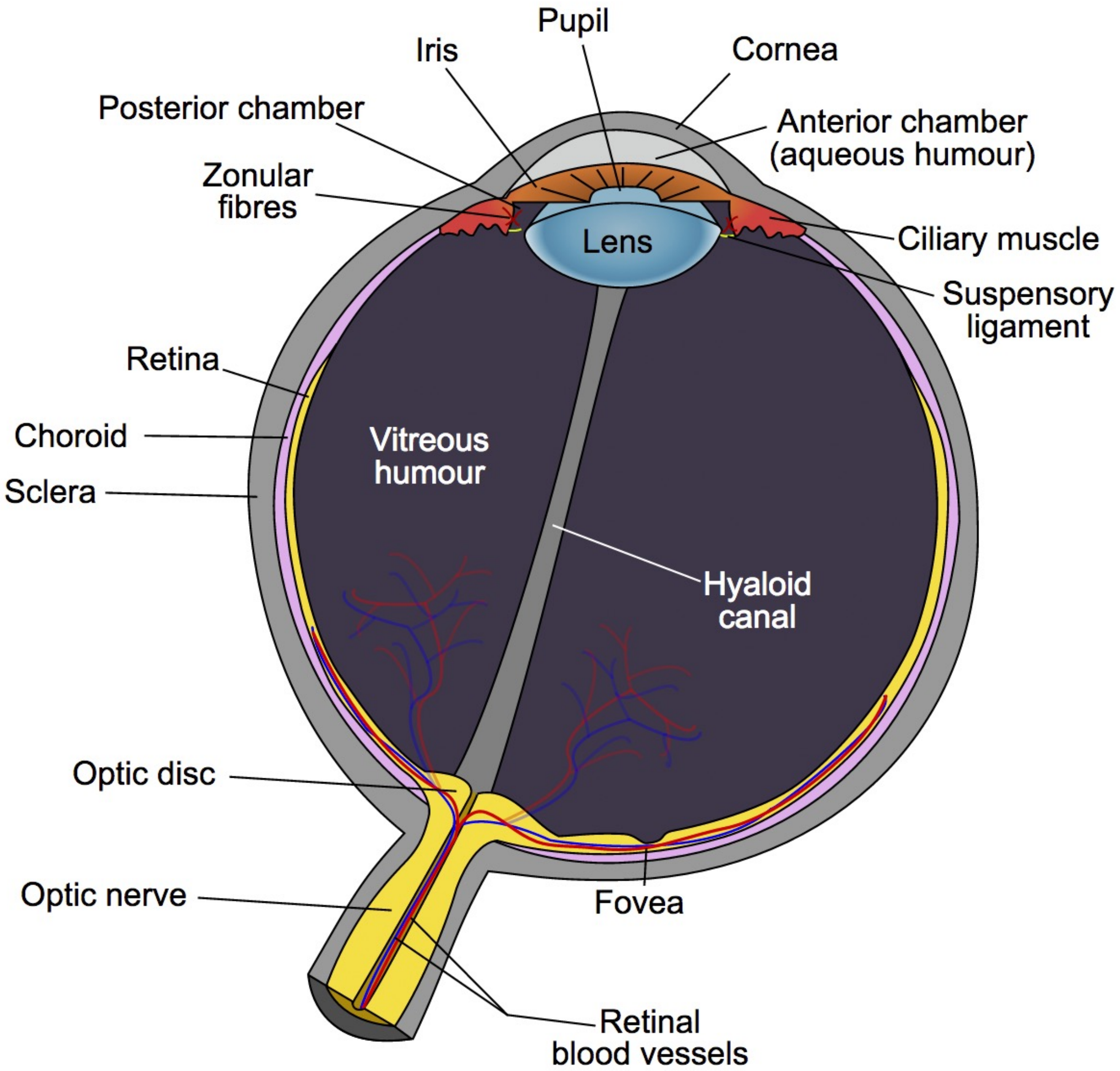


wikipedia

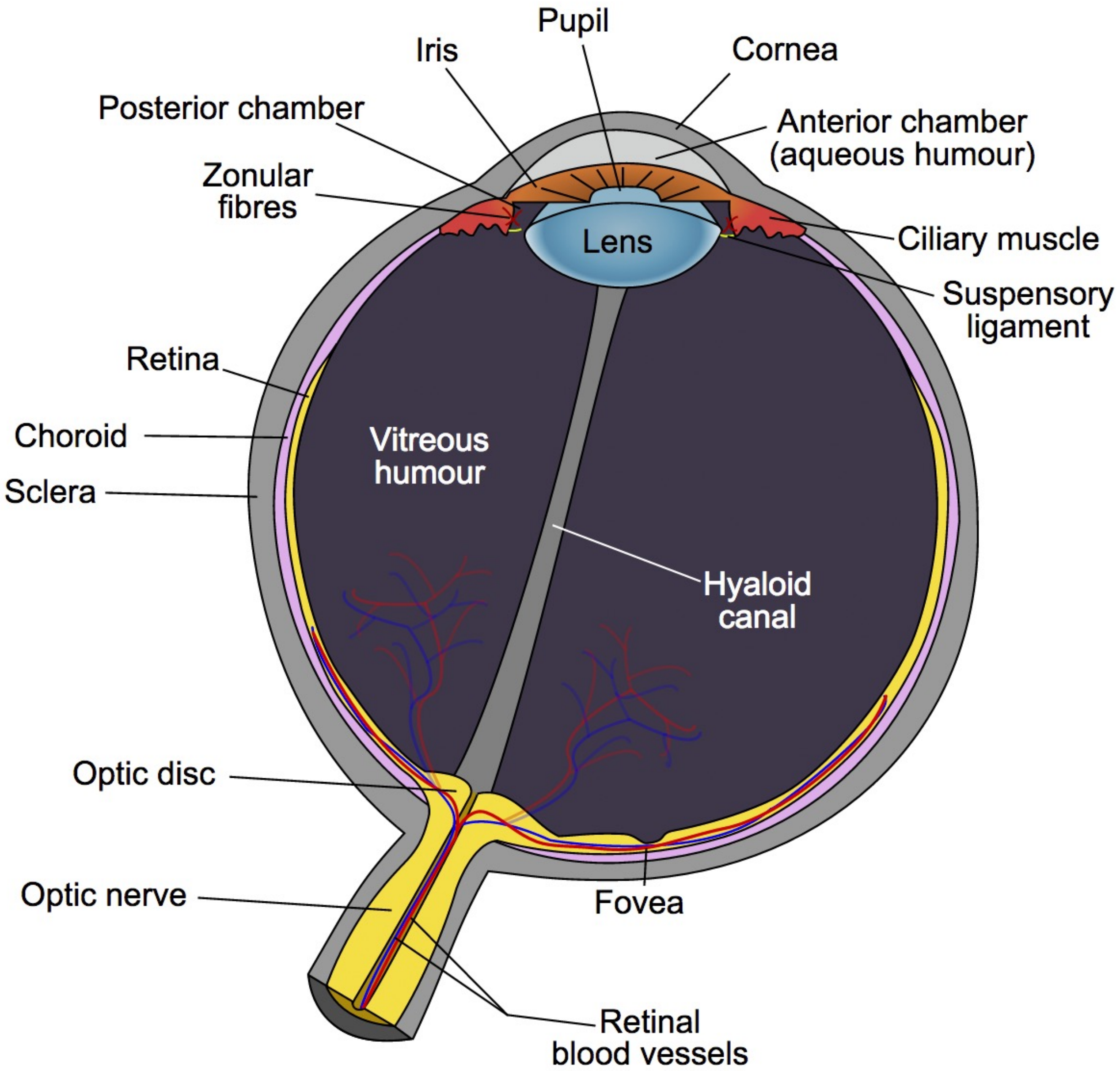
primary visual cortex

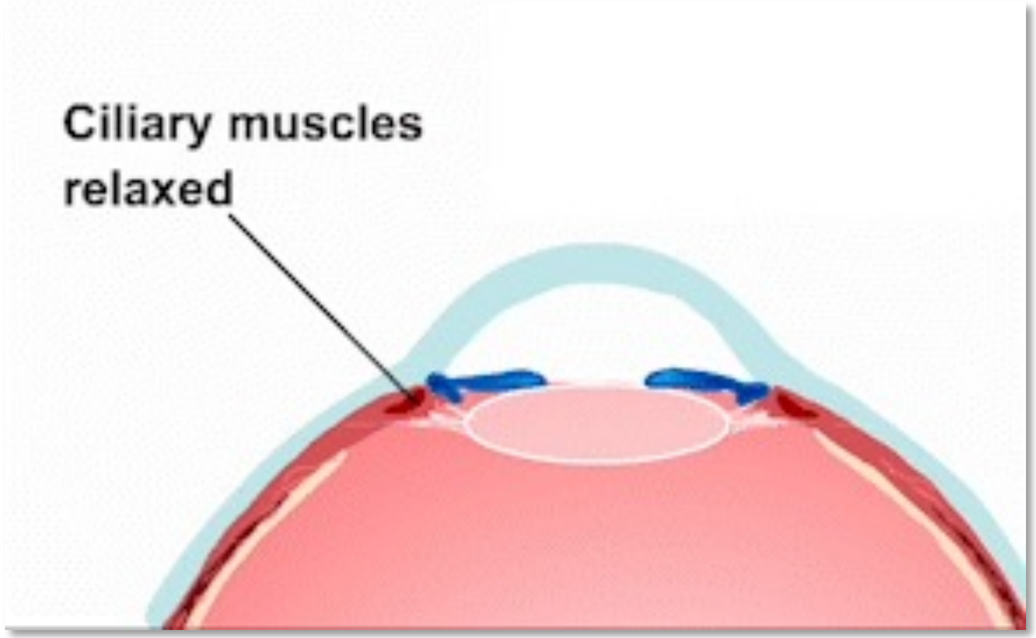
ventral stream:
recognition, object identification

Anatomy of the Human Eye

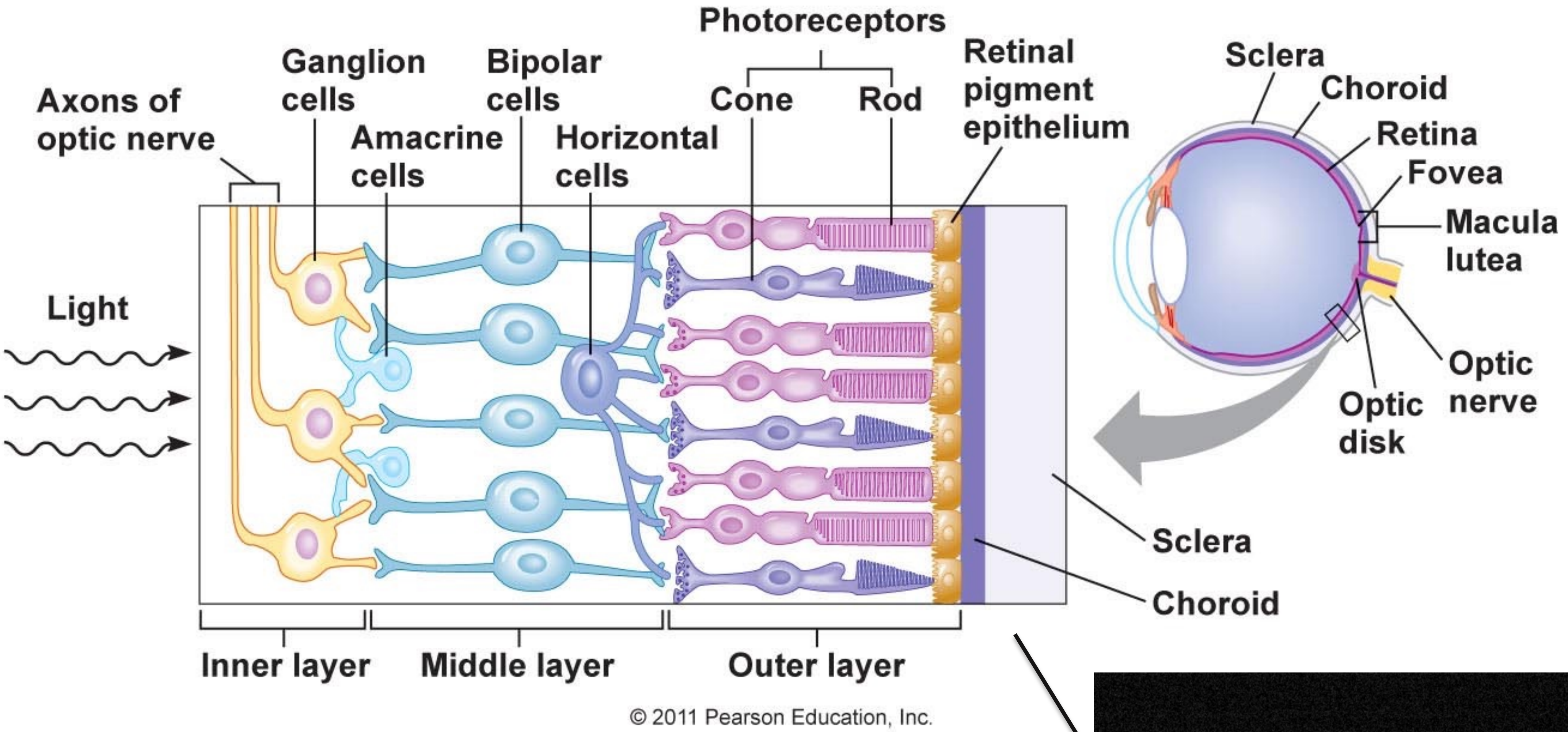


Anatomy of the Human Eye

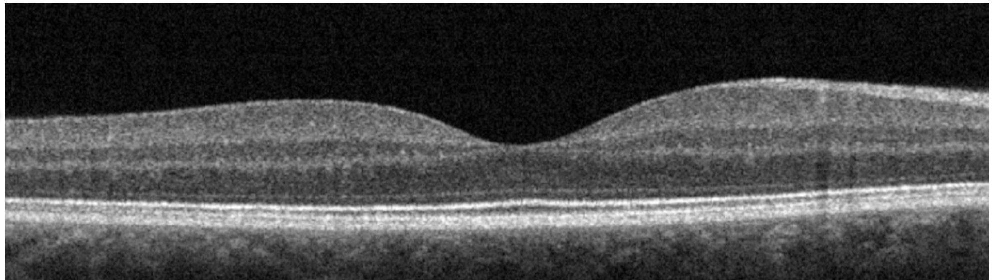




The Retina

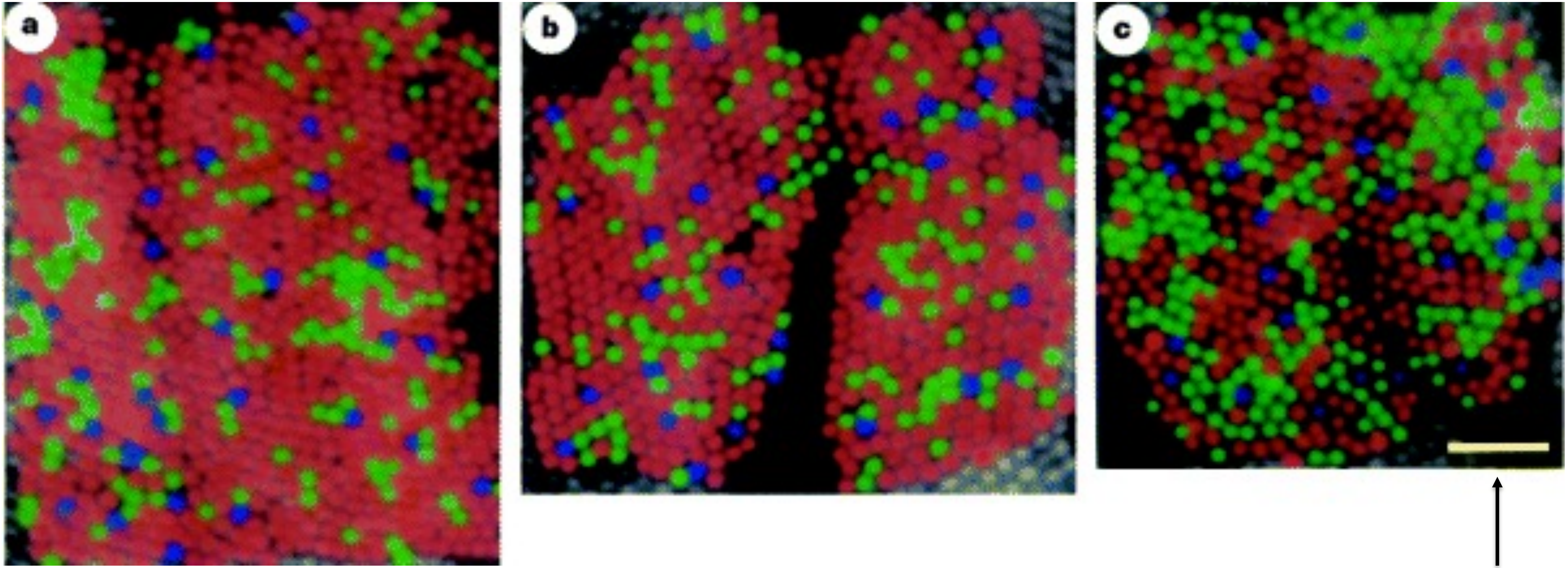


Photoreceptors
Ganglion cells
Bipolar cells
Cone
Rod
Retinal pigment epithelium
Axons of optic nerve
Amacrine cells
Horizontal cells
Light
Inner layer
Middle layer
Outer layer
Sclera
Choroid
Retina
Fovea
Macula lutea
Optic nerve
Optic disk
Sclera
Choroid
© 2011 Pearson Education, Inc.



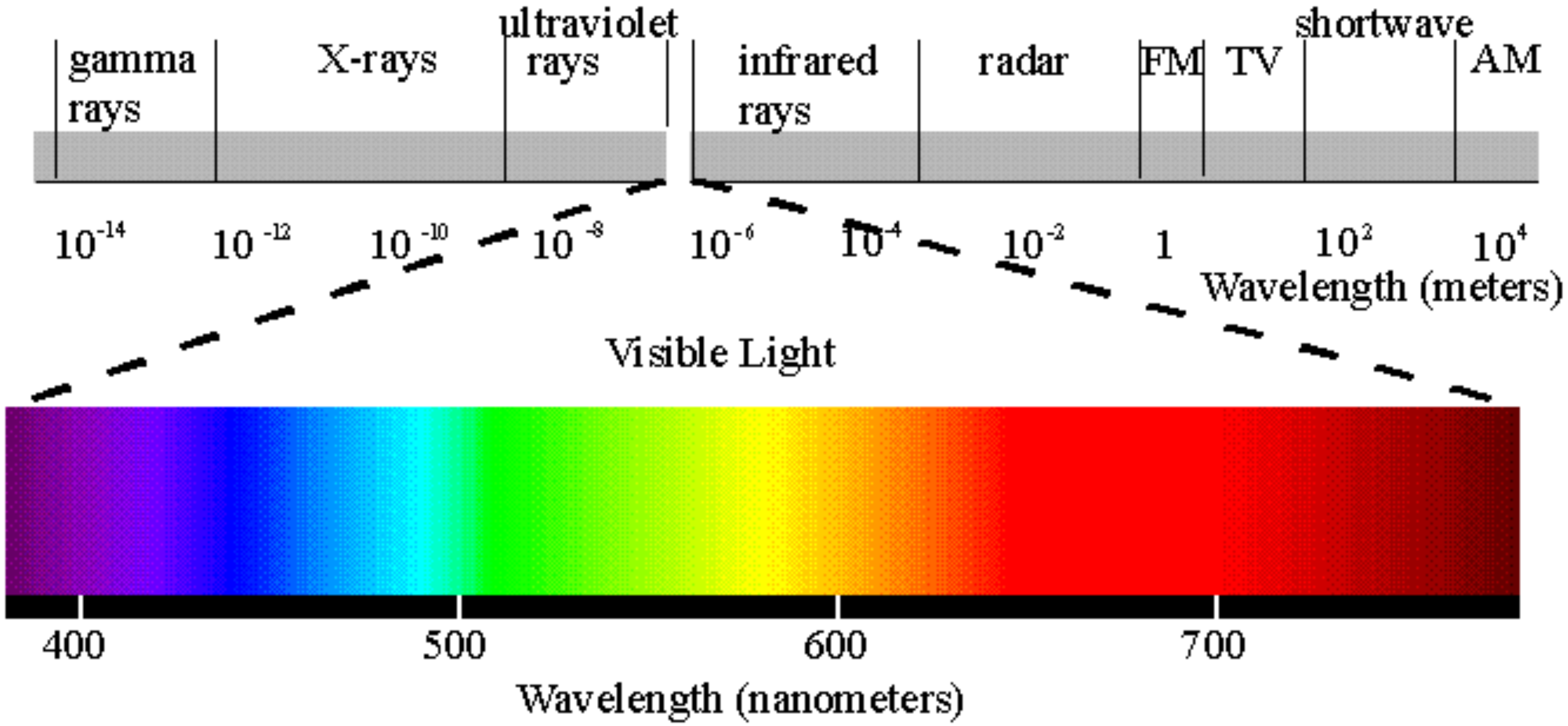
The Retina

Roorda & Williams, 1999, Nature



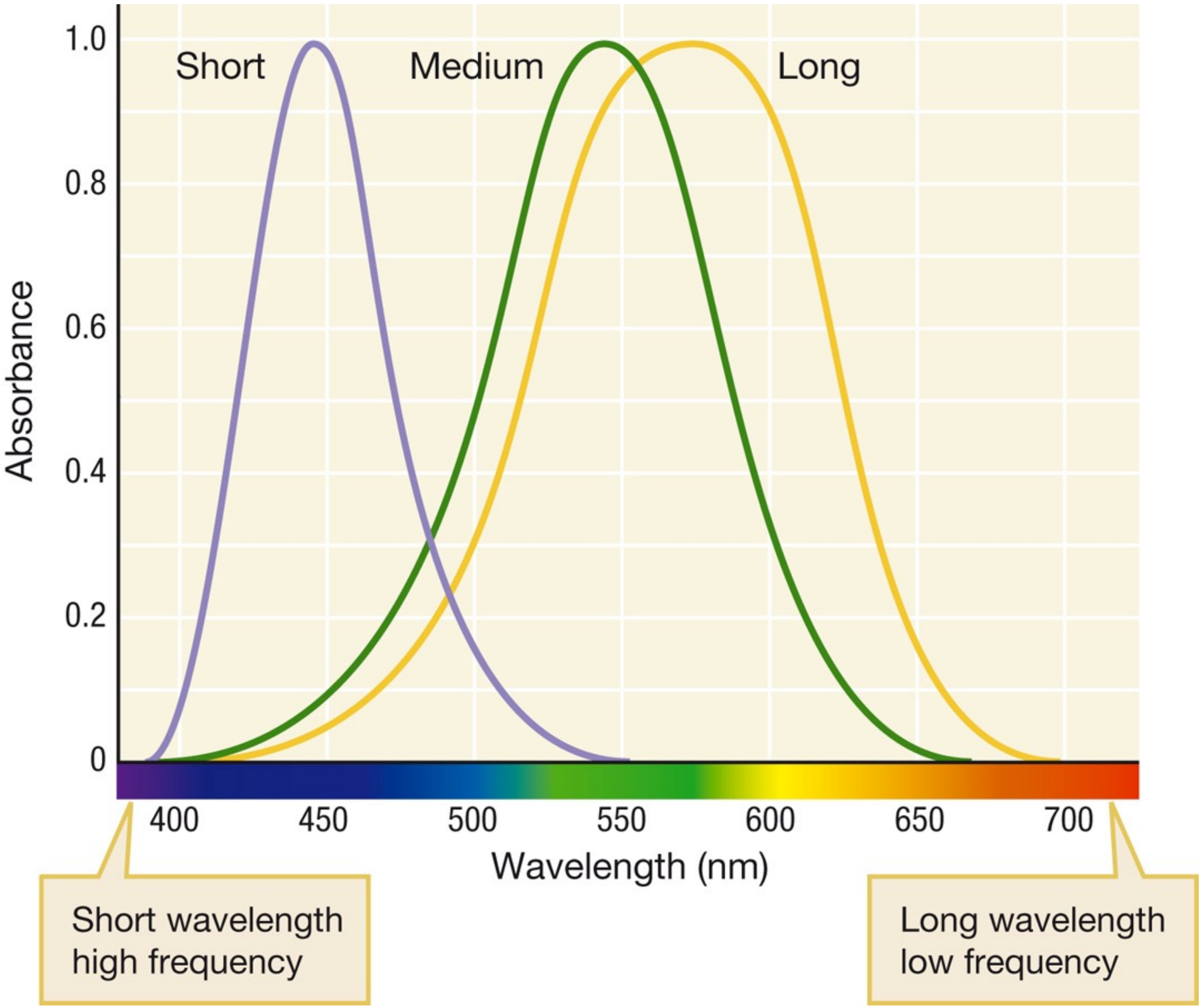
5 arcmin visual angle

Color Perception

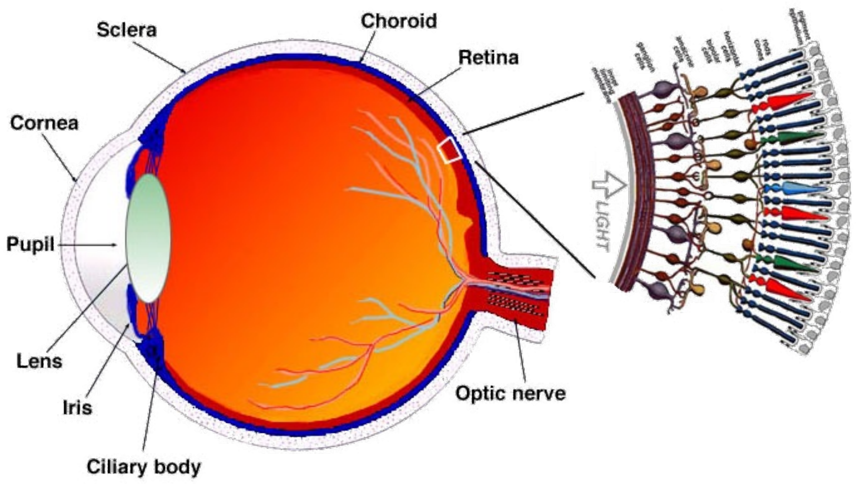


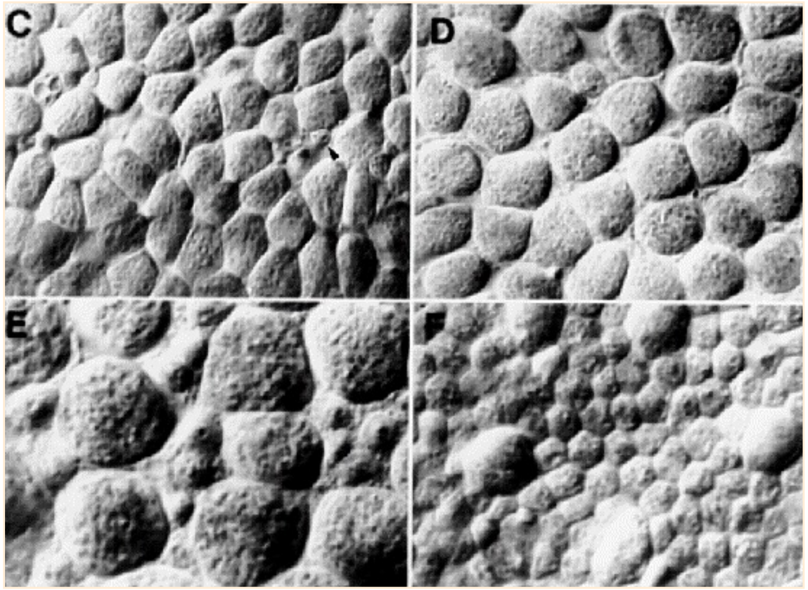
Color Perception - Sensitivity of Cones

reddit.com

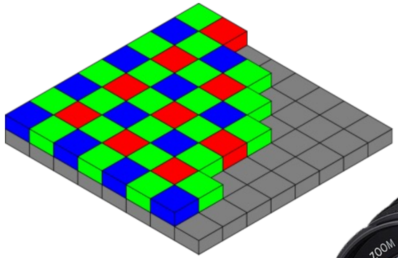


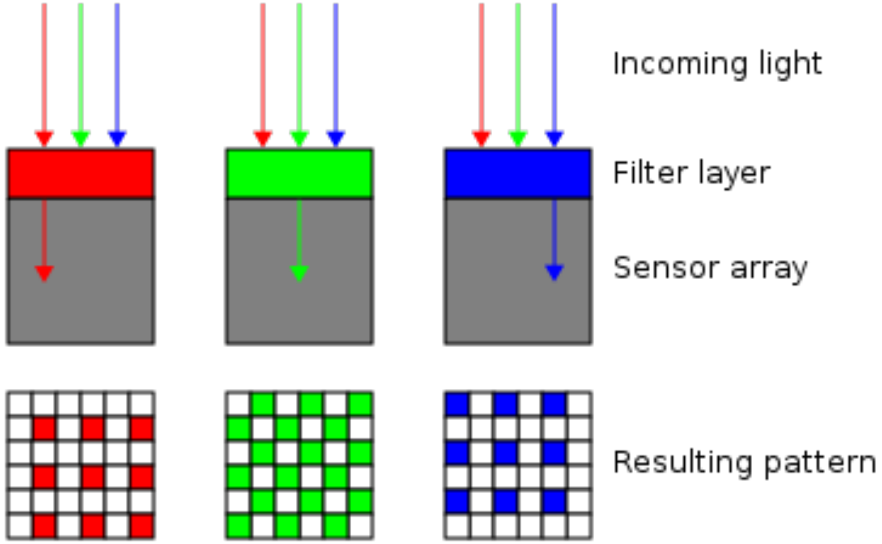
Eye vs Camera





[Williams 91]



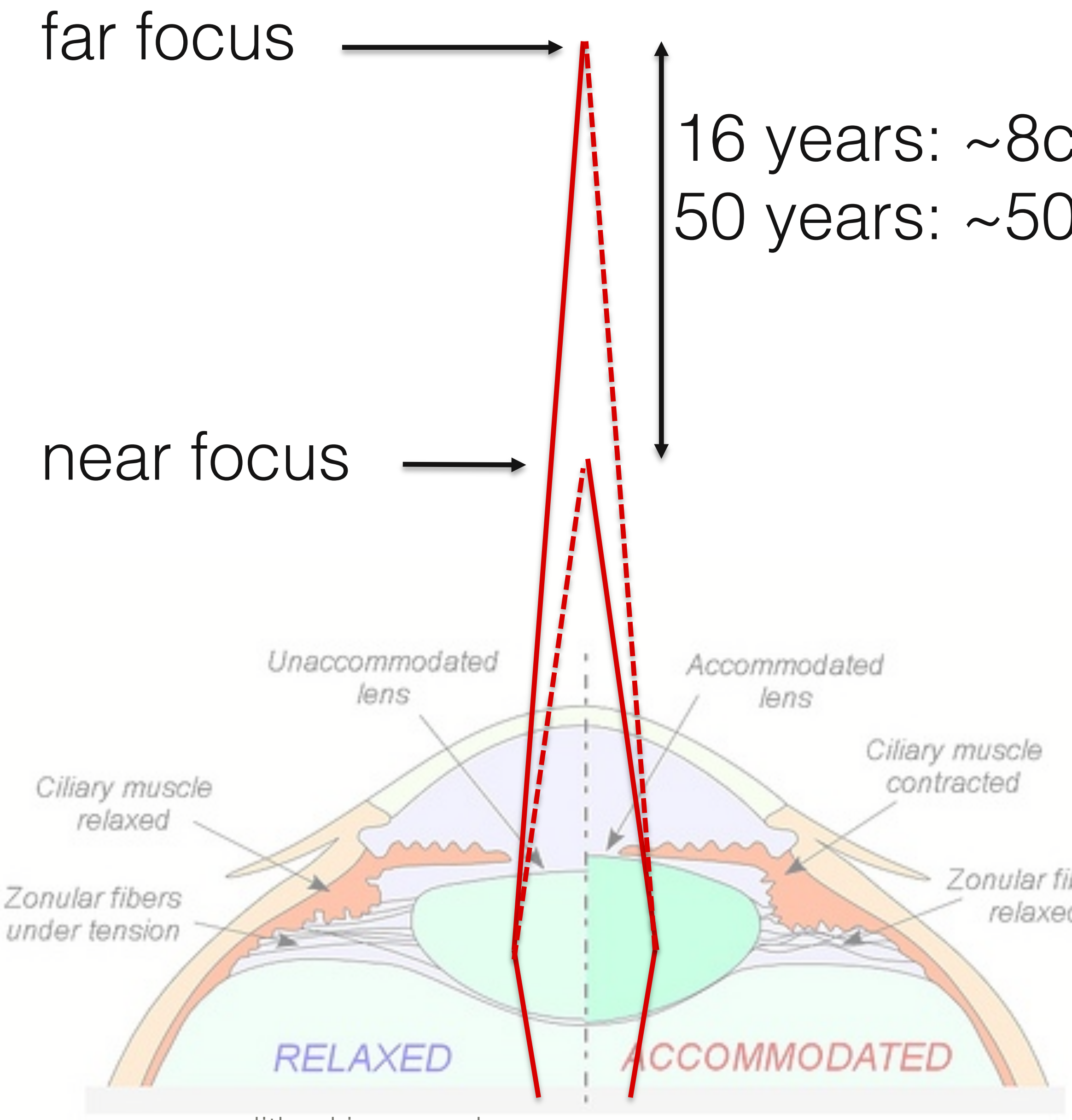


wikipedia

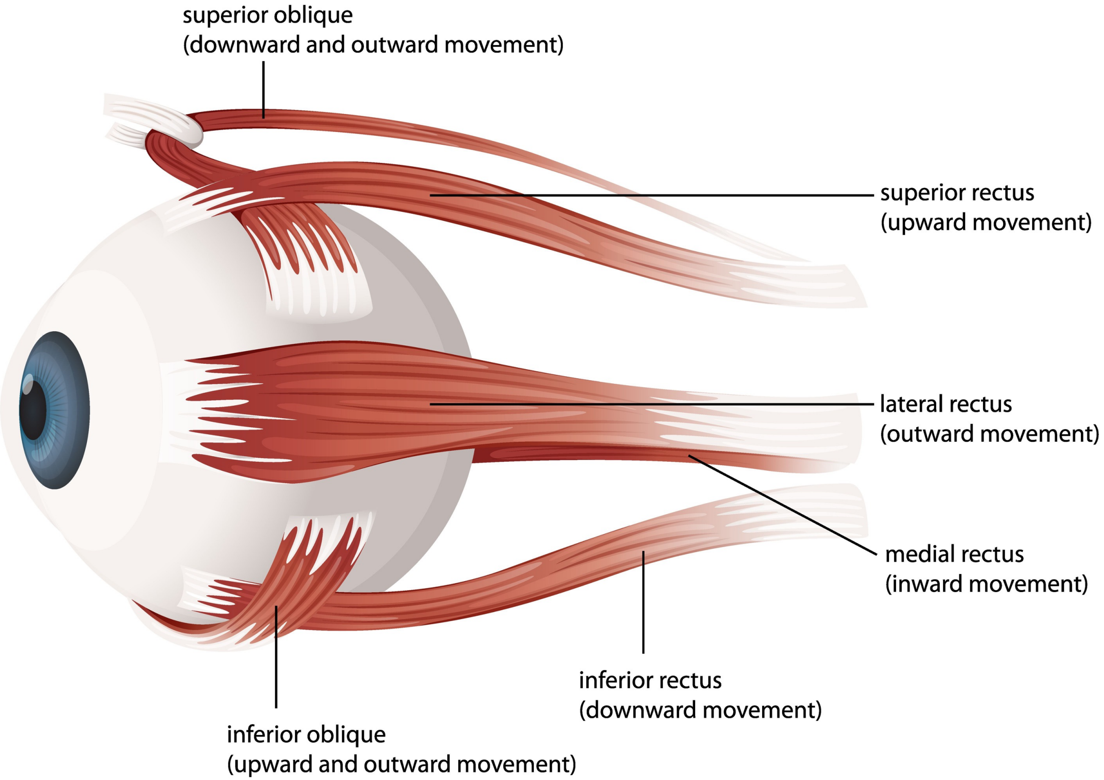
Oculomotor Processes

16 years: ~8cm to ∞

50 years: ~50cm to ∞ (mostly irrelevant)



adithyakiran.wordpress.com

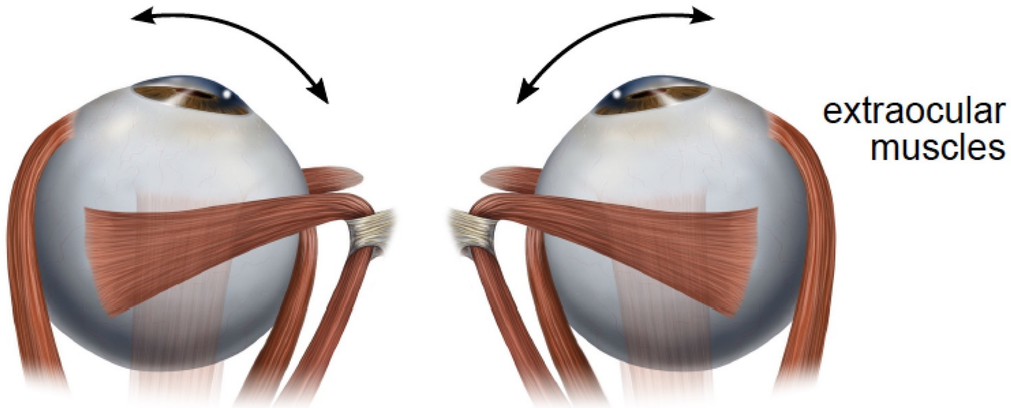


Oculumotor Processes + Visual Cues

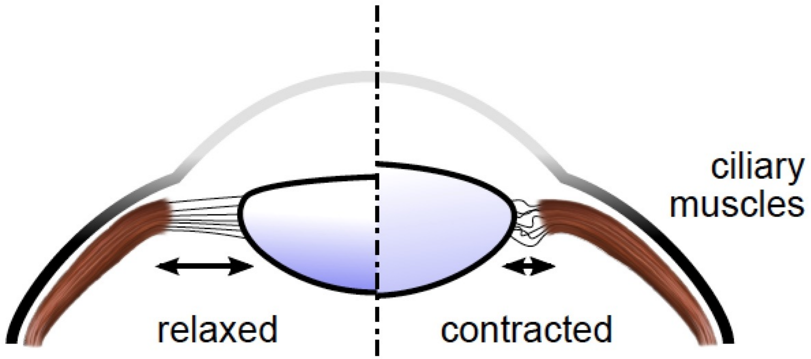
Stereopsis (Binocular)

Focus Cues (Monocular)

Oculomotor Cue



Vergence

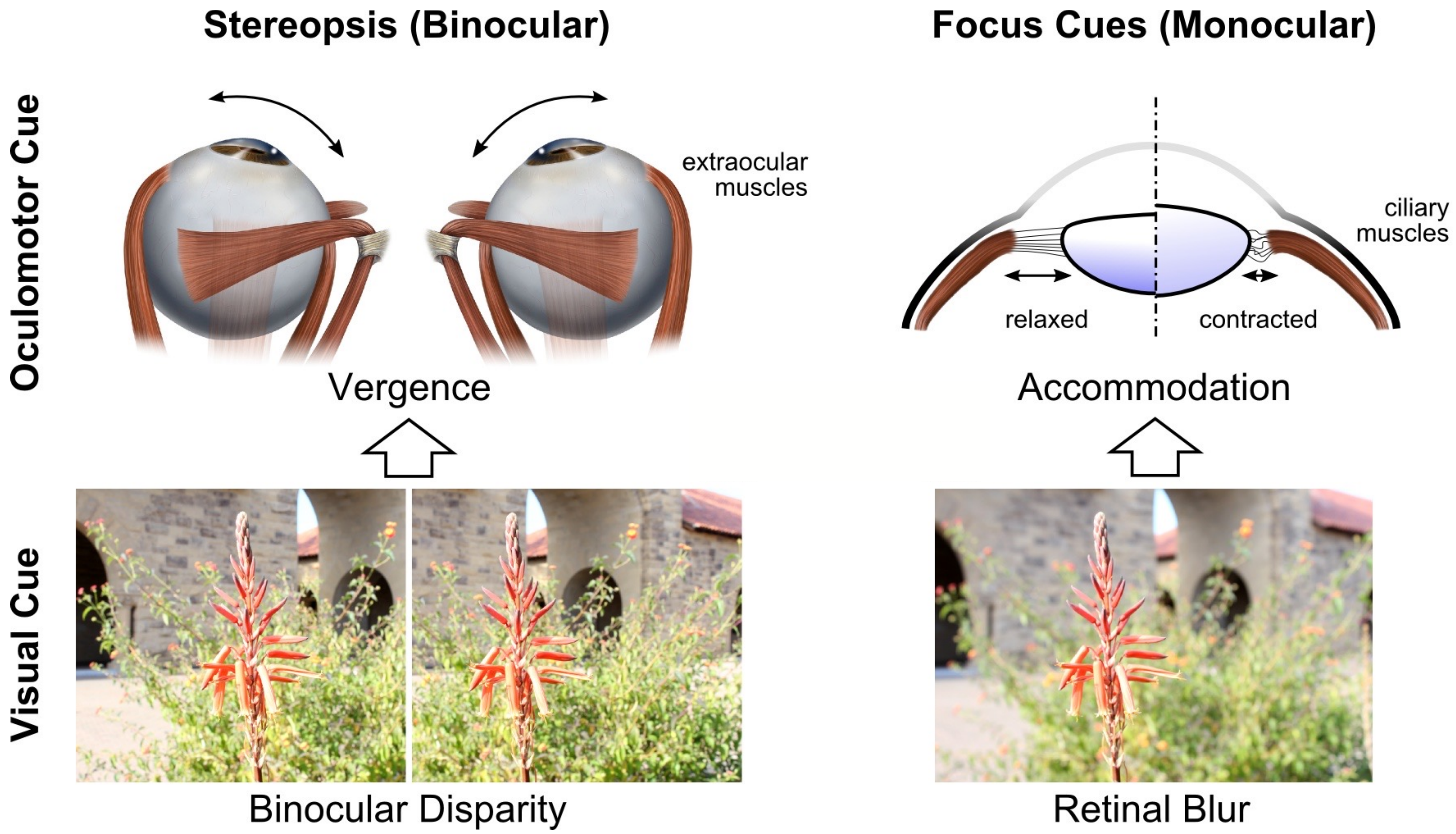


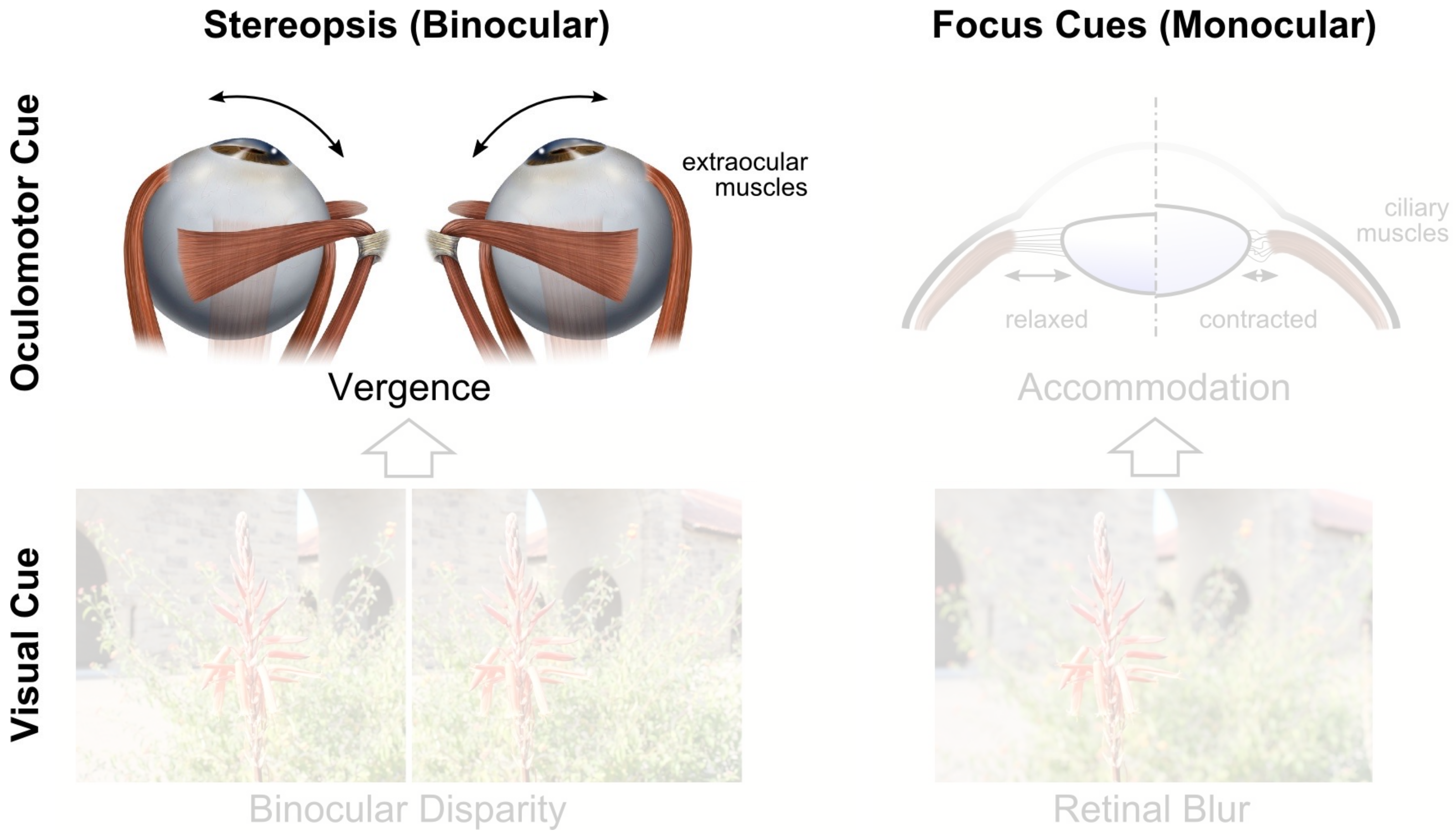
Accommodation

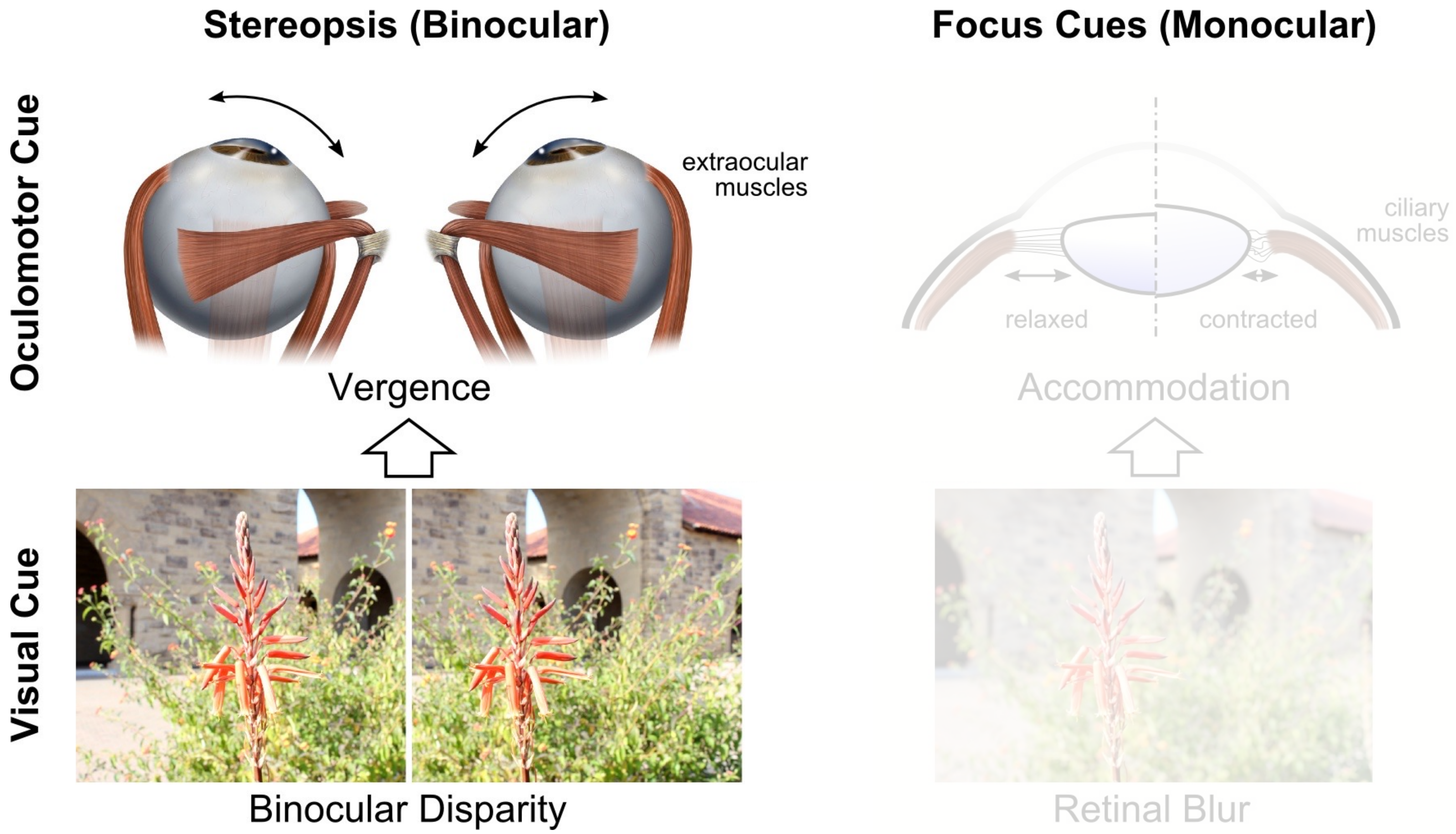
Visual Cue

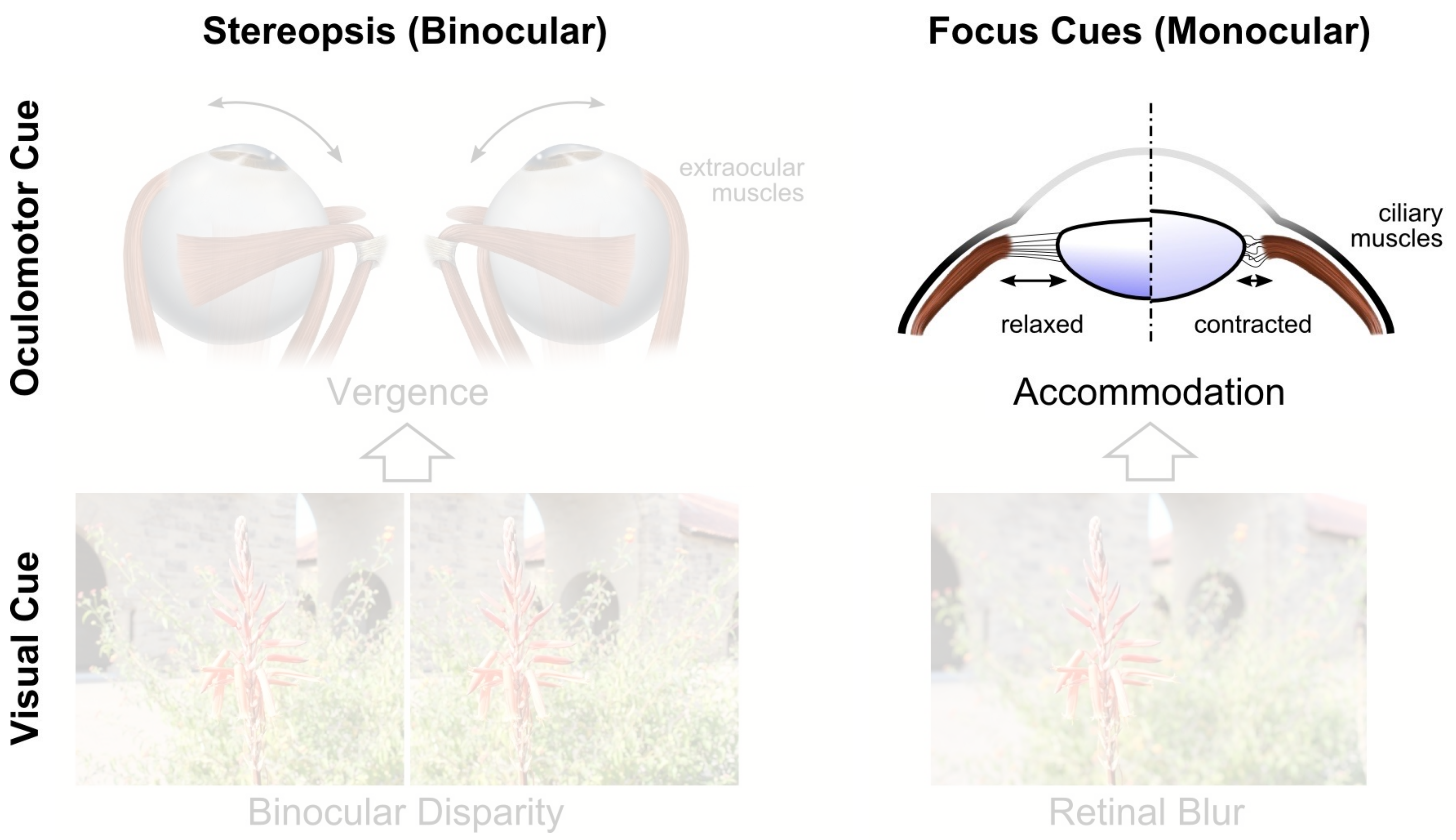
Binocular Disparity

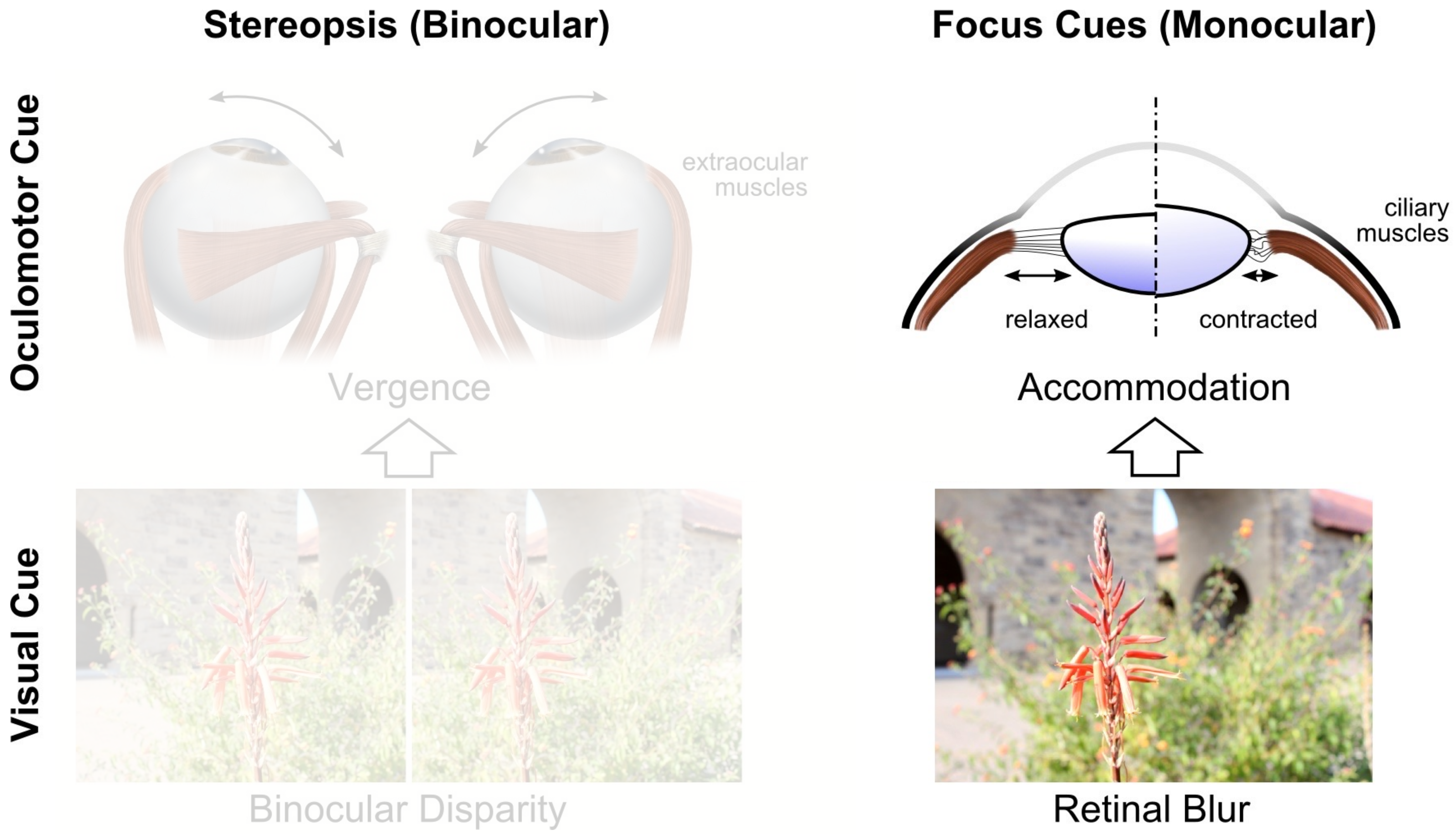
Retinal Blur

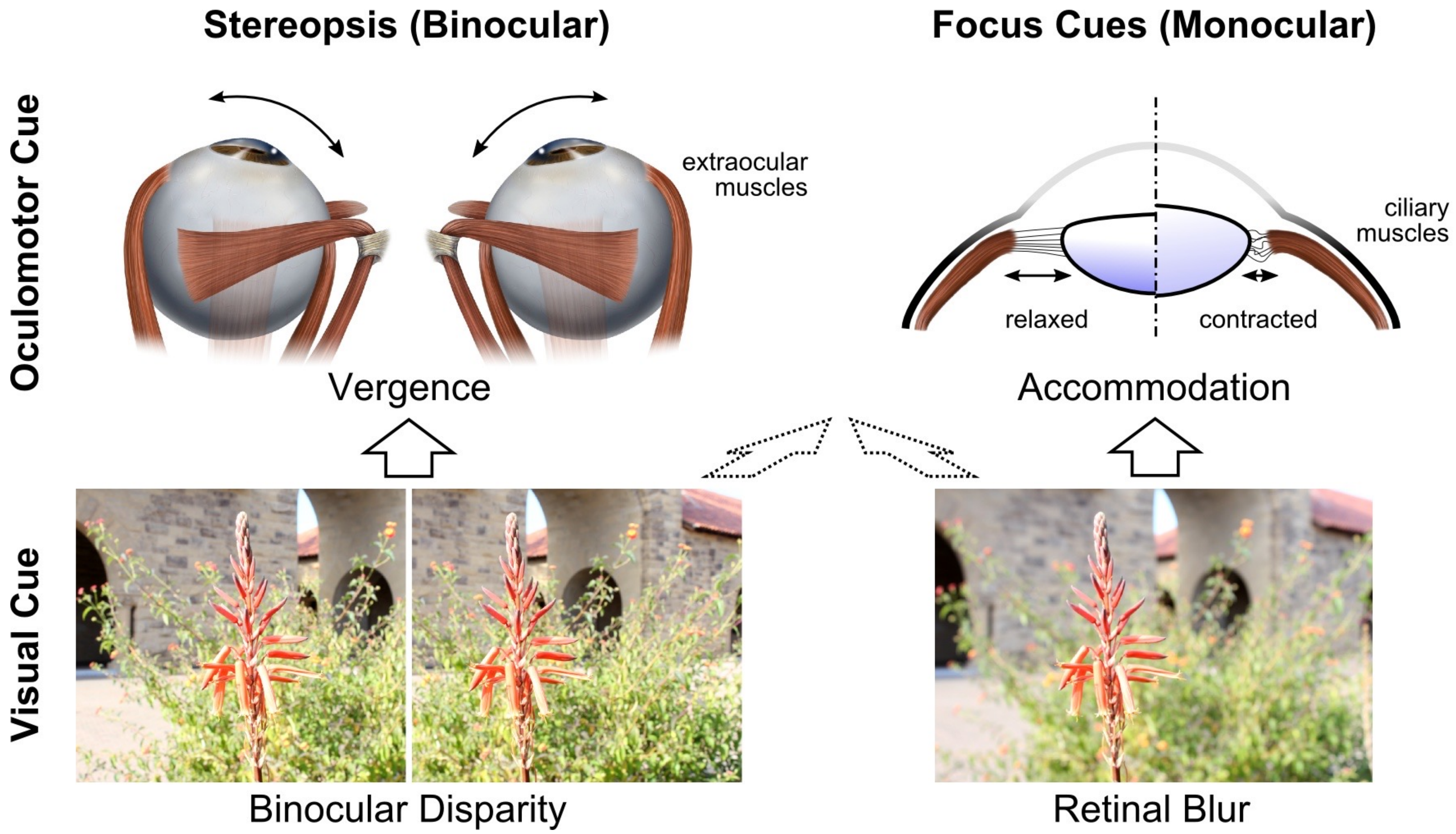

Stereopsis (Binocular)
Focus Cues (Monocular)
Oculomotor Cue
extraocular muscles
ciliary muscles
relaxed
contracted
Vergence
Accommodation
Visual Cue
Binocular Disparity
Retinal Blur


Stereopsis (Binocular)
Focus Cues (Monocular)
Oculomotor Cue
Visual Cue
extraocular muscles
Vergence
ciliary muscles
relaxed
contracted
Accommodation
Binocular Disparity
Retinal Blur

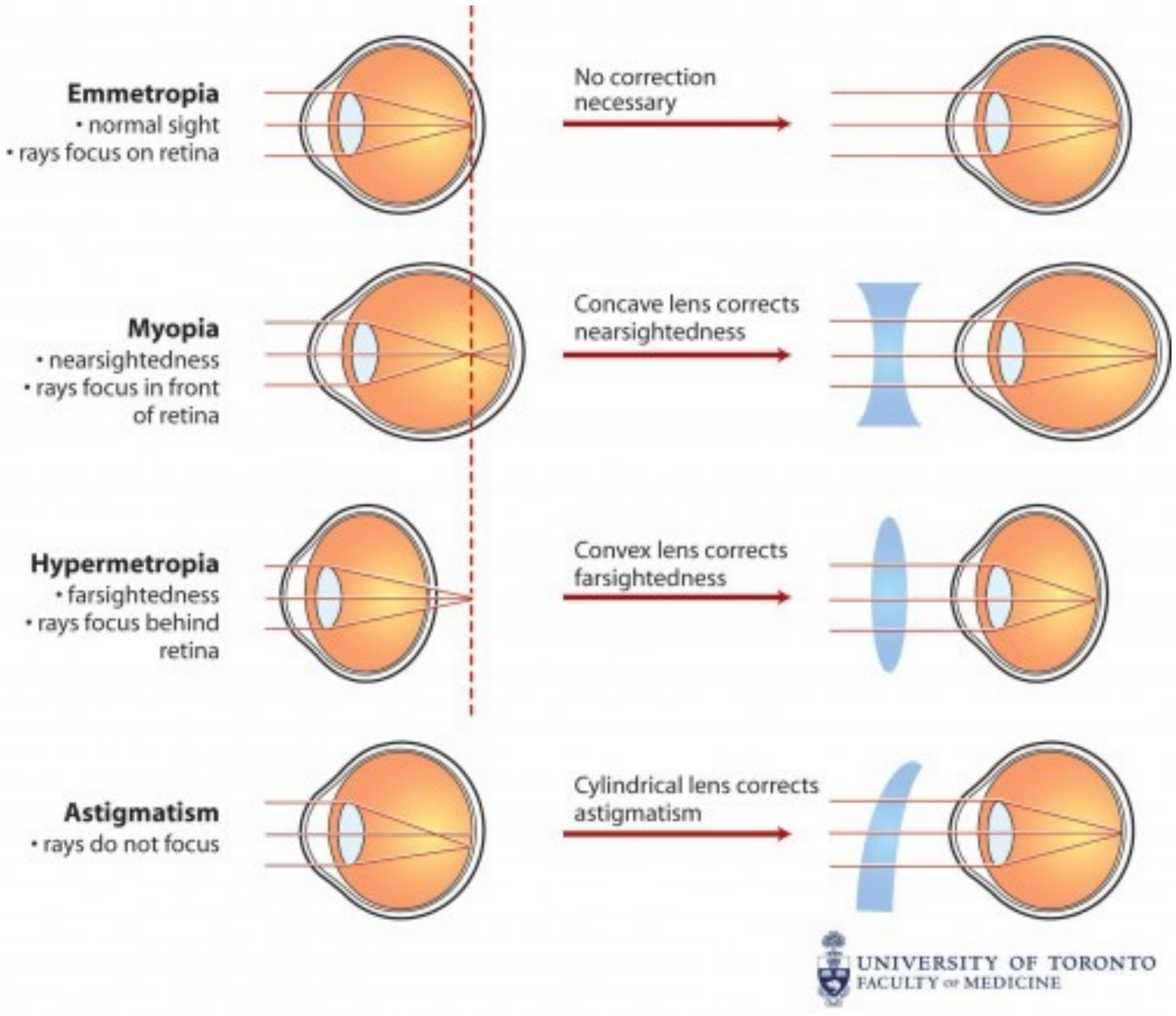

Stereopsis (Binocular)
Focus Cues (Monocular)
Oculomotor Cue
Visual Cue
extraocular
muscles
Vergence
Binocular Disparity
ciliary
muscles
relaxed
contracted
Accommodation
Retinal Blur


Stereopsis (Binocular)
Focus Cues (Monocular)
Oculomotor Cue
Visual Cue
extraocular
muscles
Vergence
ciliary
muscles
relaxed
contracted
Accommodation
Binocular Disparity
Retinal Blur


Stereopsis (Binocular)
Focus Cues (Monocular)
Oculomotor Cue
Visual Cue
extraocular muscles
Vergence
ciliary muscles
relaxed
contracted
Accommodation
Binocular Disparity
Retinal Blur

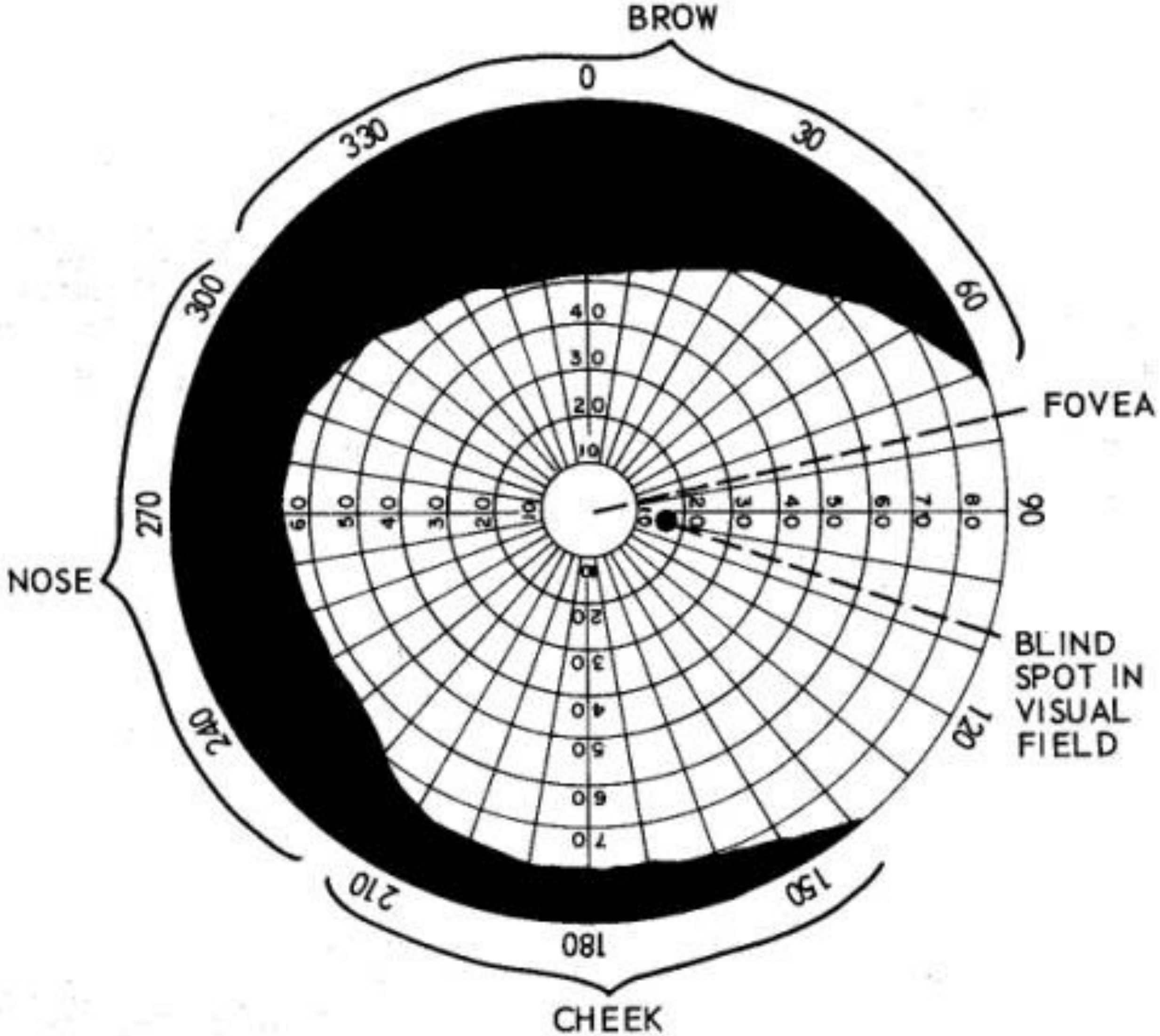

Stereopsis (Binocular)
Focus Cues (Monocular)
Oculomotor Cue
extraocular muscles
ciliary muscles
relaxed
contracted
Vergence
Accommodation
Visual Cue
Binocular Disparity
Retinal Blur

Refractive Errors

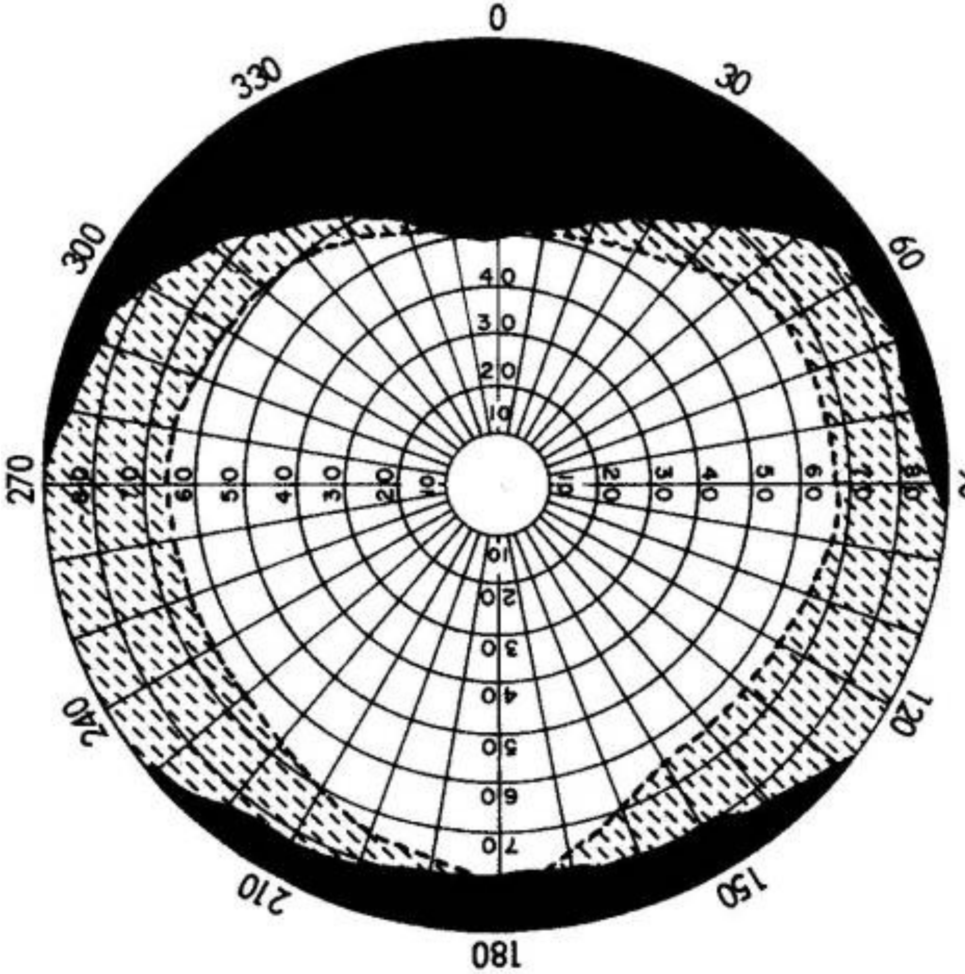


Visual Field / Field of View

Ruch & Fulton, 1960

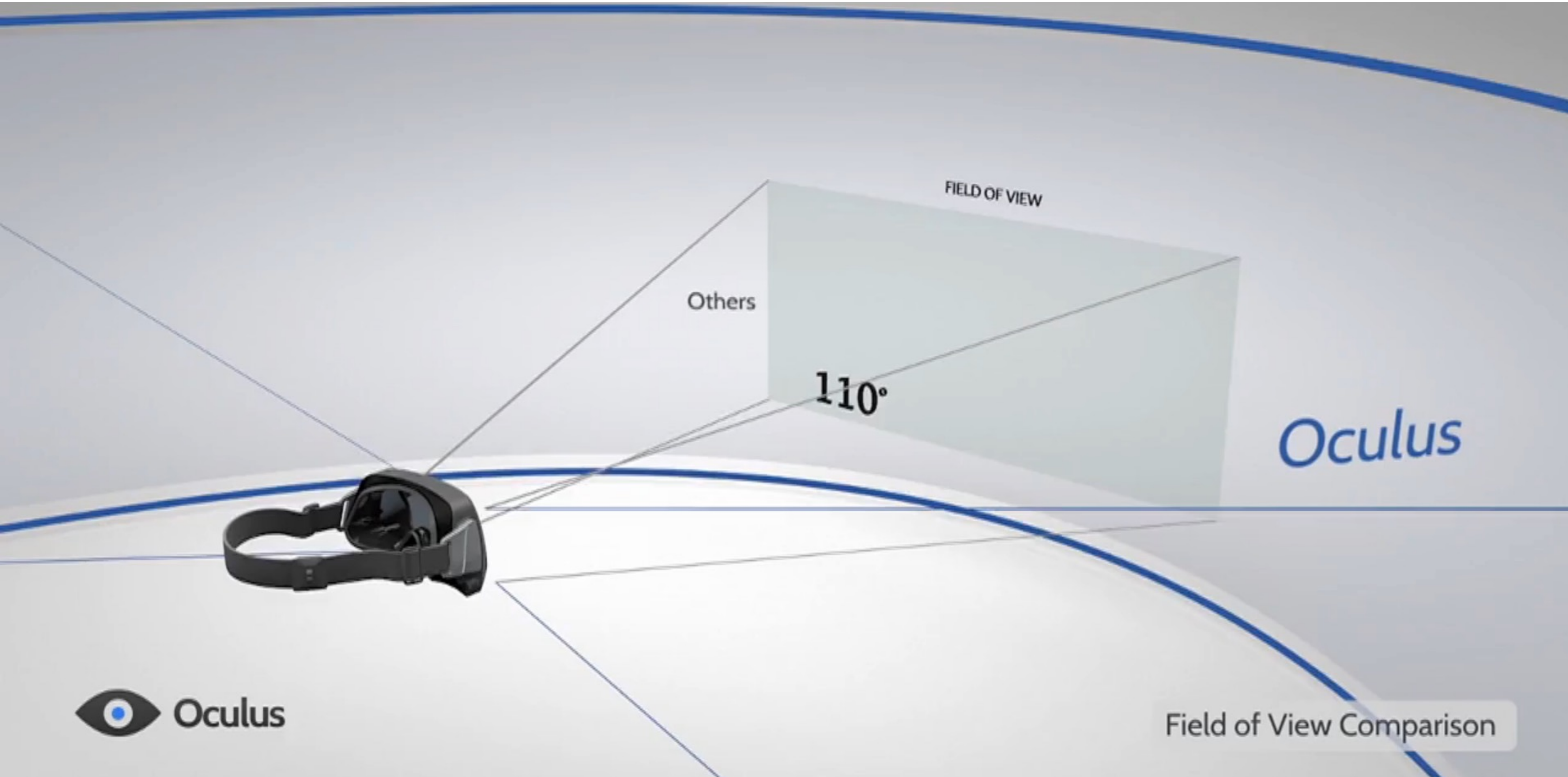


monocular visual field



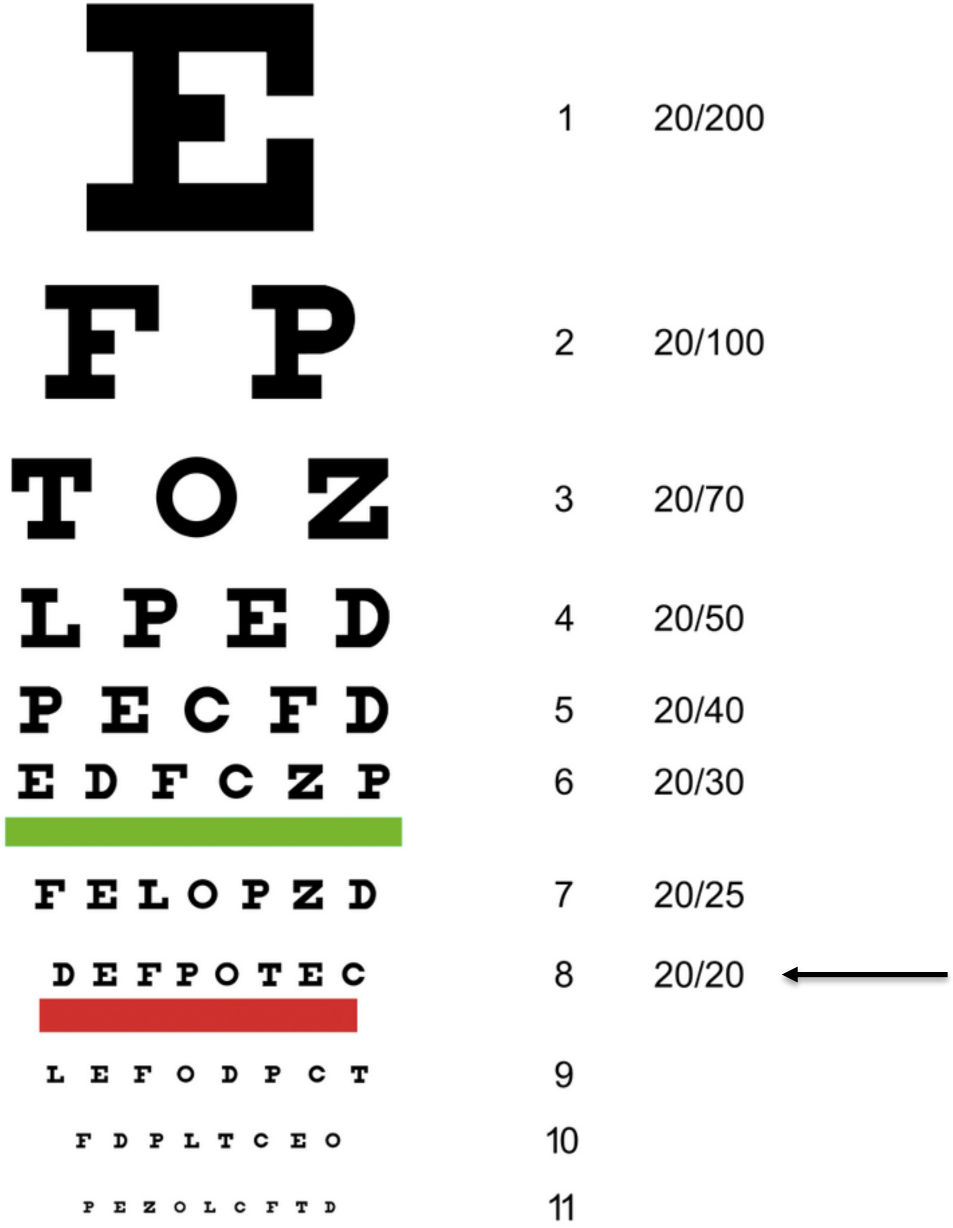
binocular visual field

Immersive VR – How Important is the FOV?



Visual Acuity

Snellen chart



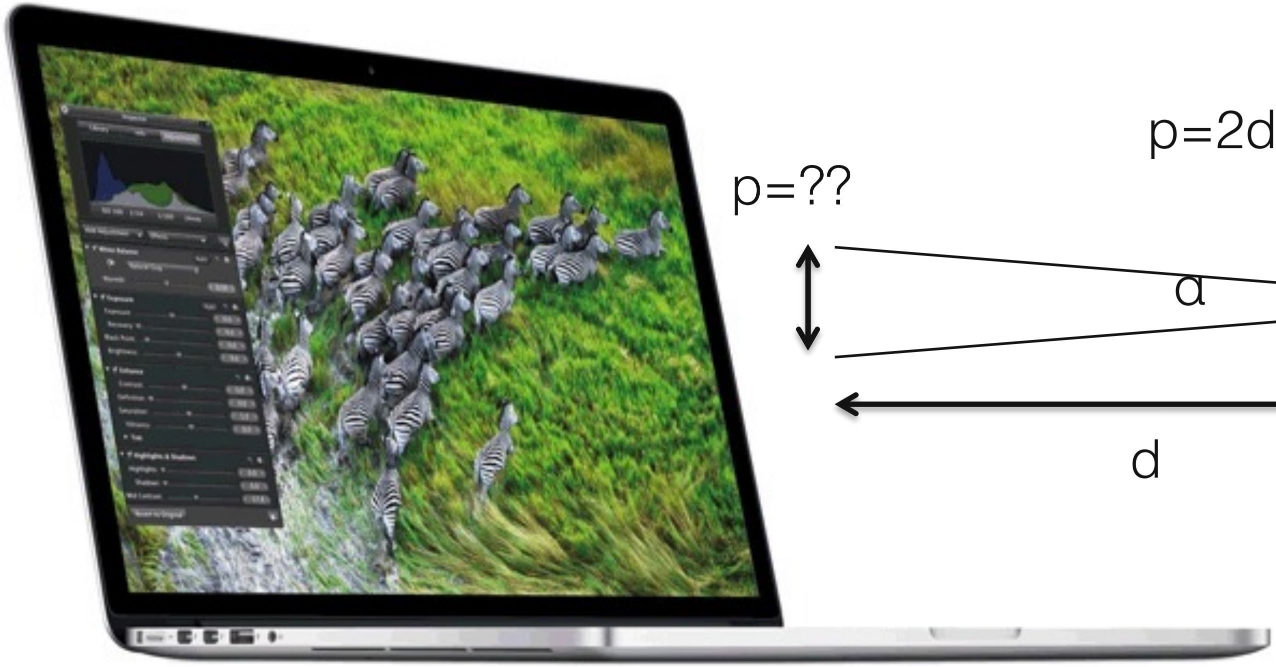
characters are 5 arc min, need to resolve 1 arc min to read

Retina Displays

Steve Jobs: 300 dpi is retina resolution

our math: ~286 dpi

Apple Inc.



$p = 2d \tan(\alpha/2)$

p=??

α

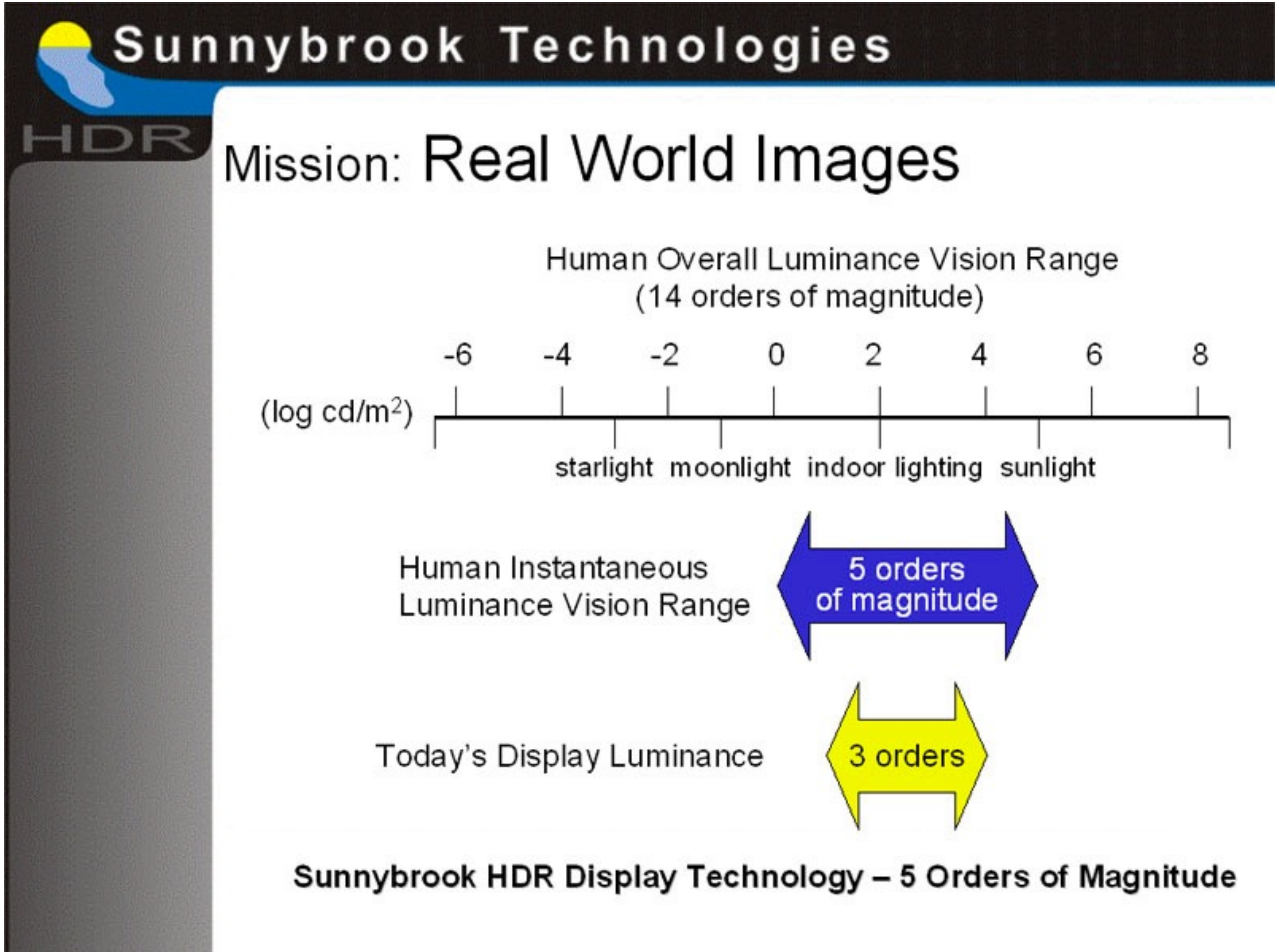
eye

d

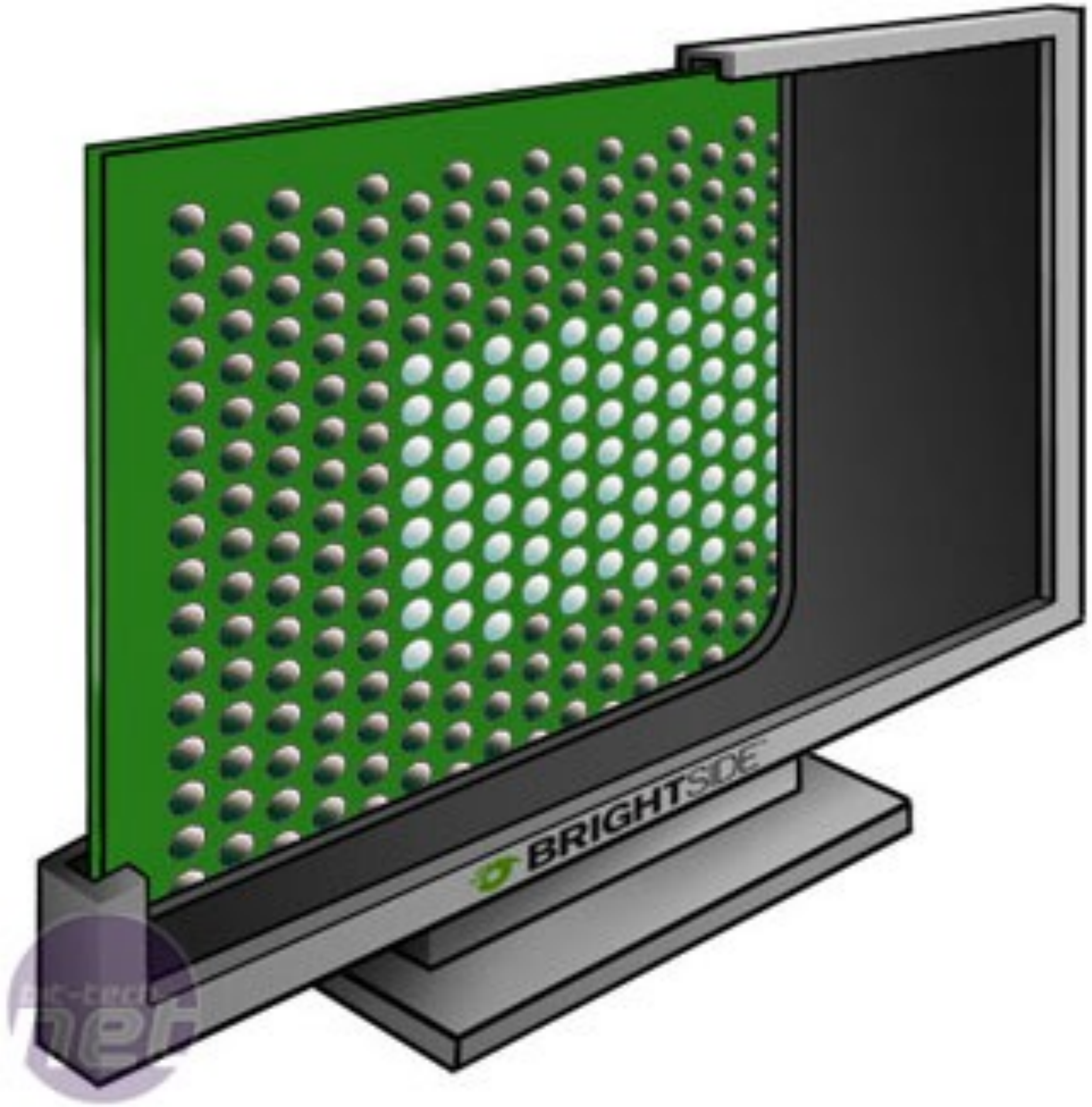
tablet, 12” away, resolvable pixel:

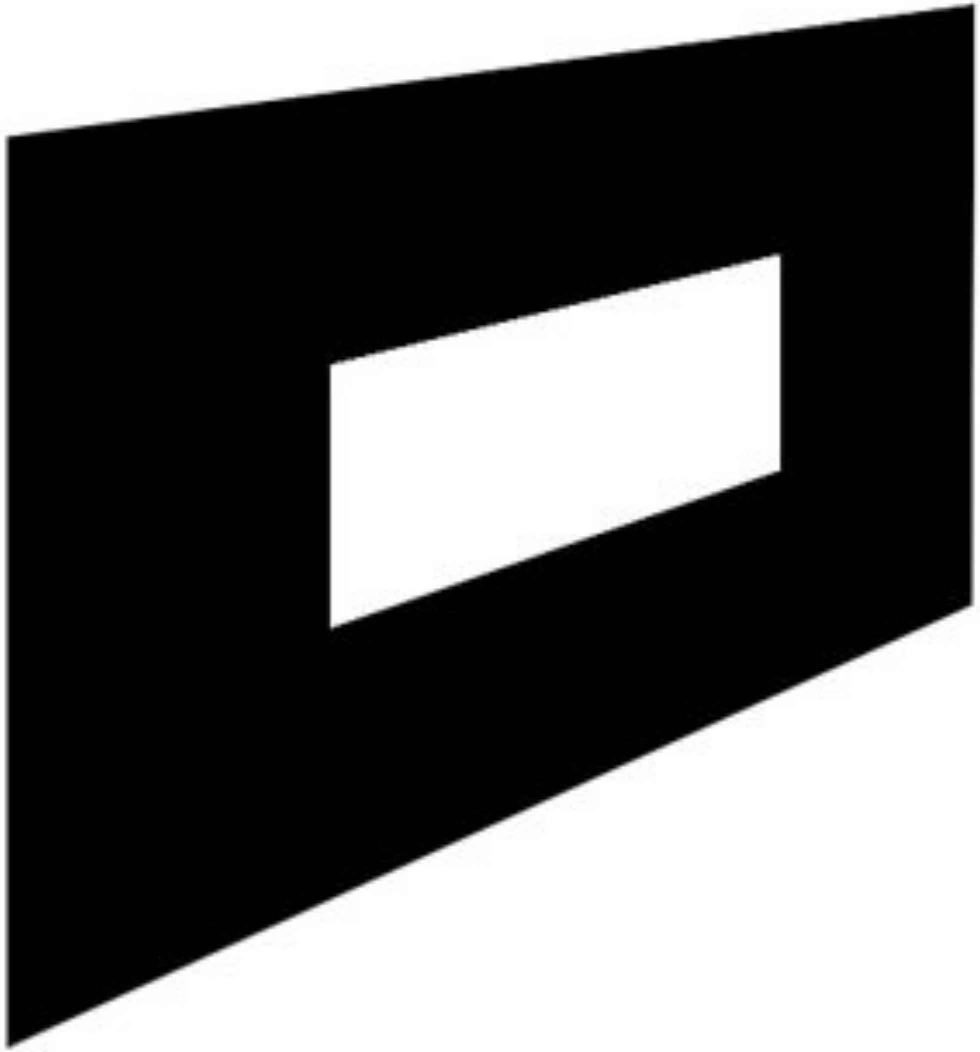
$p = 2 \cdot 12'' \cdot \tan(1 \text{ arc min}/2) = 0.0035''$

Dynamic Range



High Dynamic Range Displays

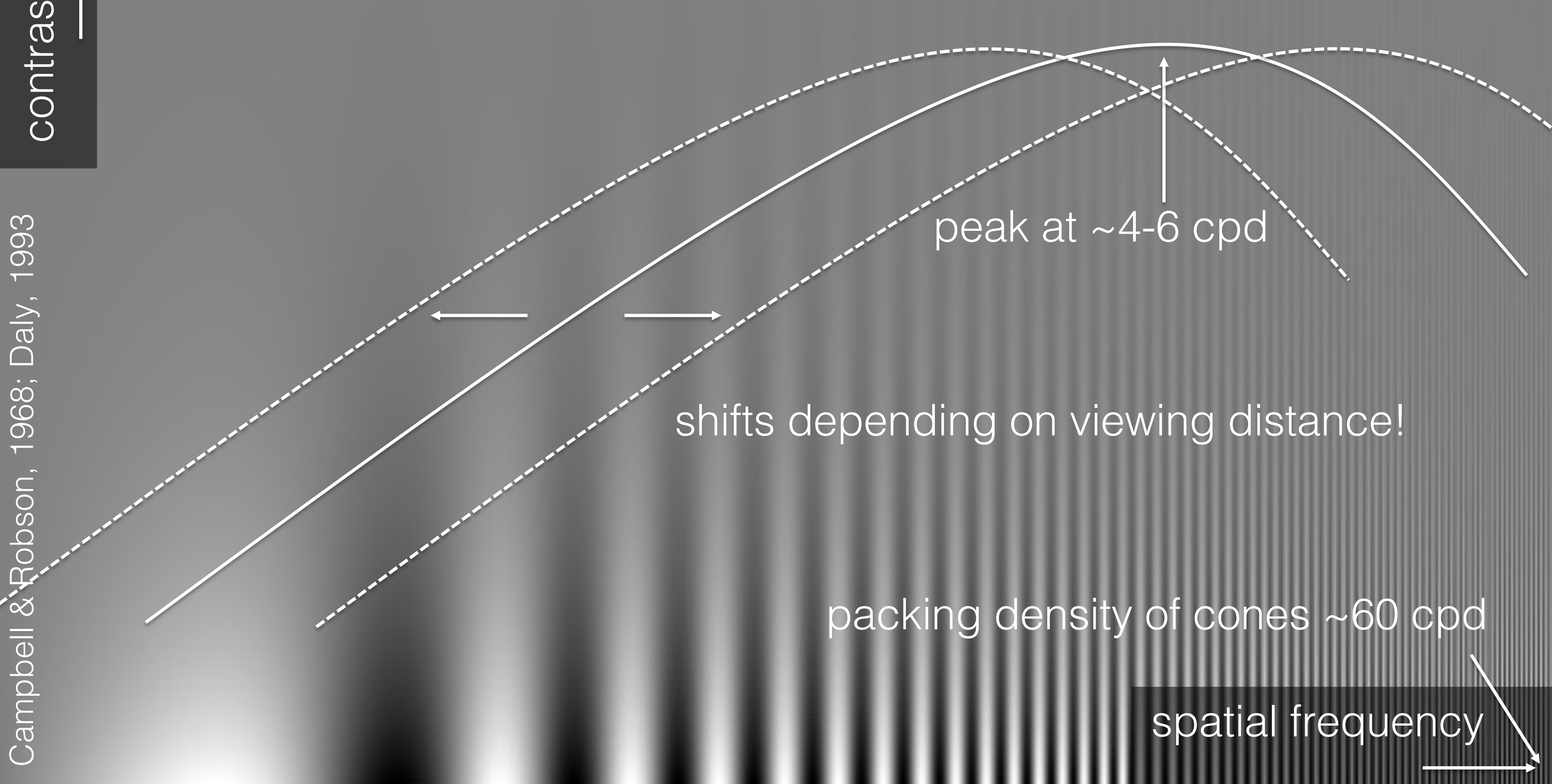




Contrast

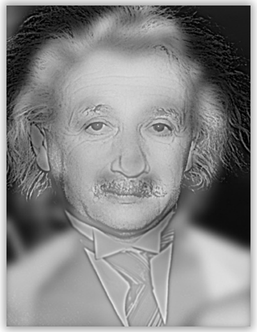
Which image has a higher contrast?

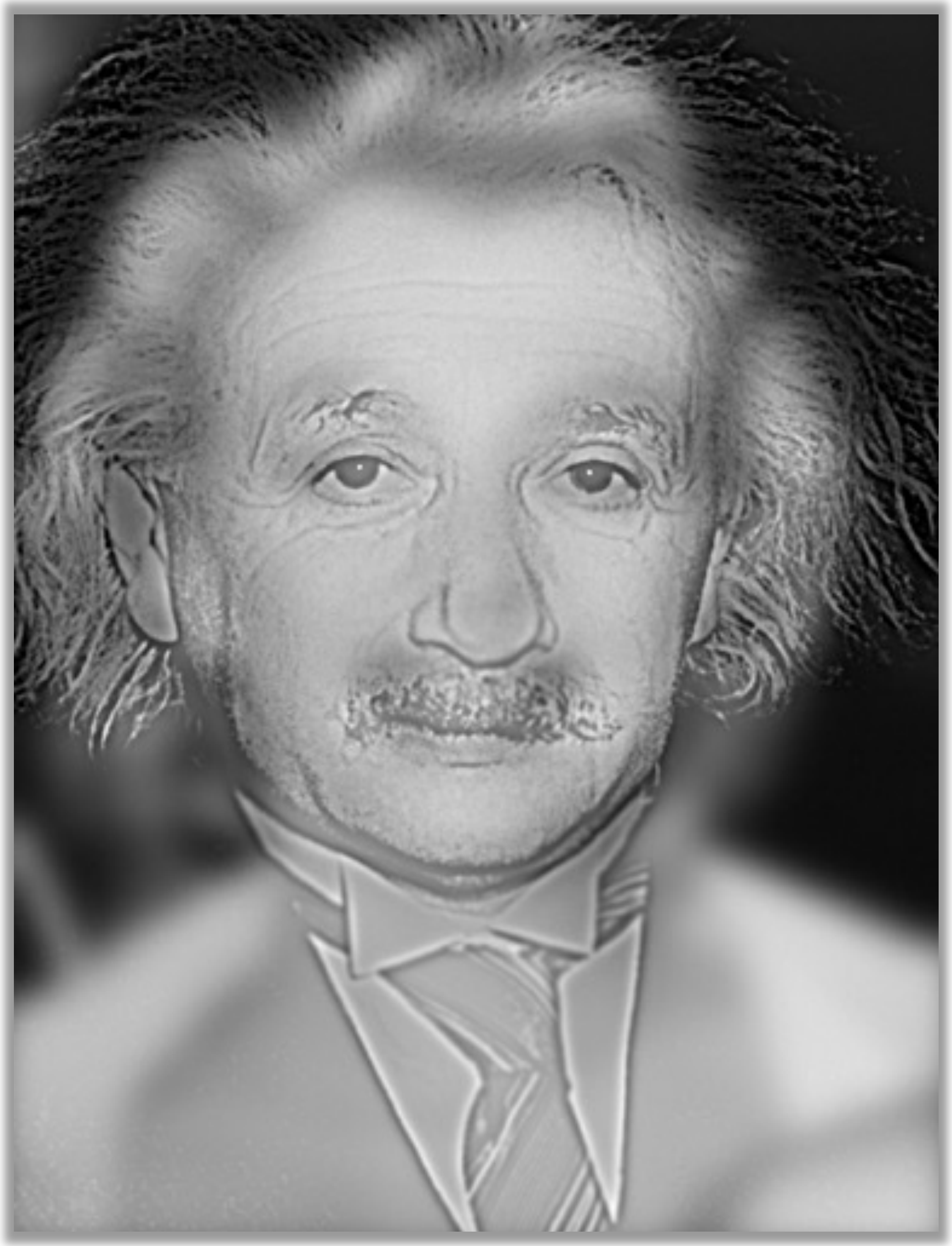
Contrast Sensitivity Function



Hybrid Images

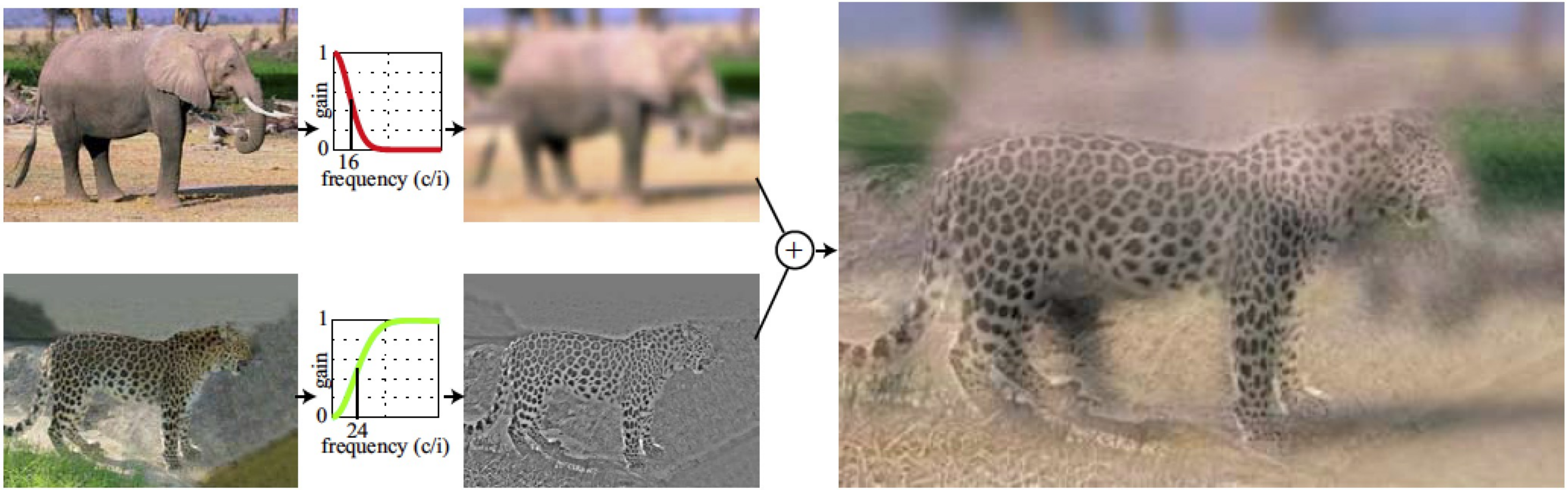
Oliva, Torralba, & Schyns, 2006, ACM SIGGRAPH





Hybrid Images

Oliva, Torralba, & Schyns, 2006, ACM SIGGRAPH



Depth Perception

wikipedia

Depth Perception

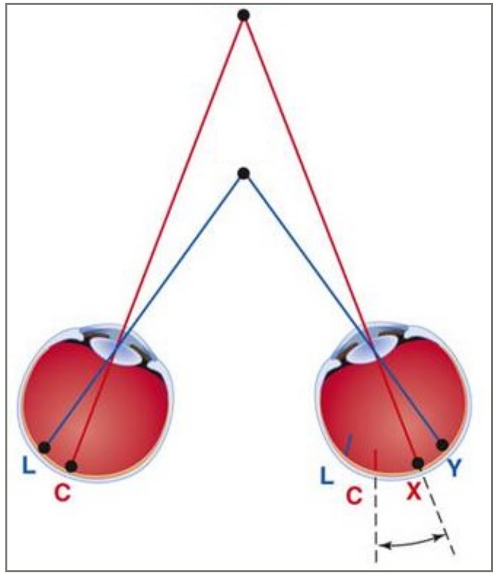
monocular cues

- perspective
- relative object size
- absolute size
- occlusion
- accommodation
- retinal blur
- motion parallax
- texture gradients
- shading
- …

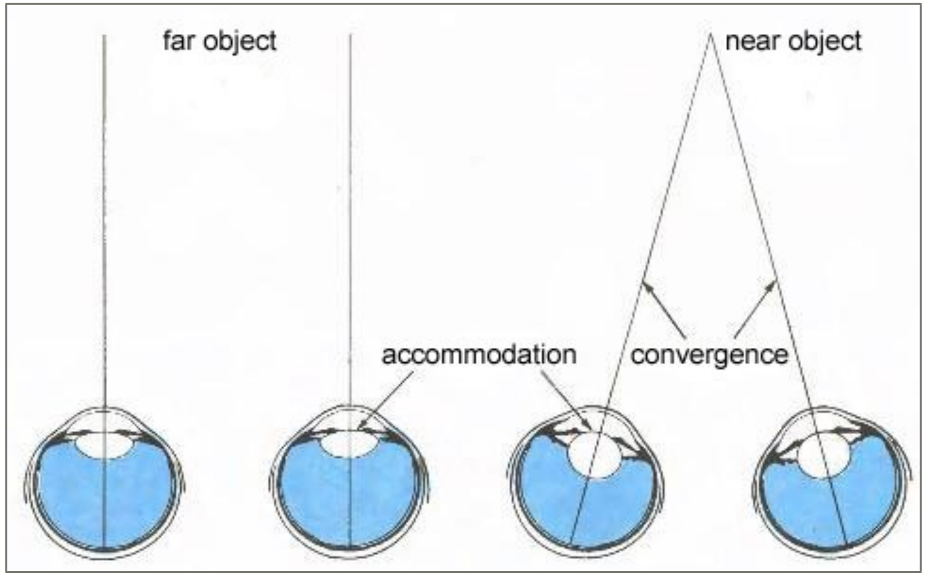
binocular cues

- (con)vergence
- disparity / parallax
- …

Depth Perception

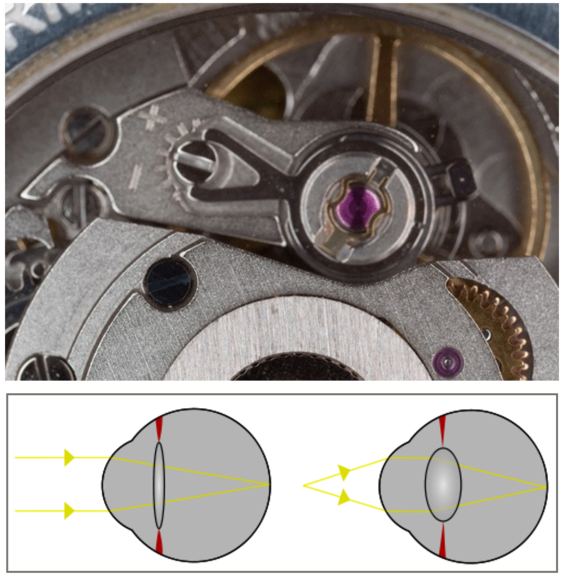


binocular disparity



convergence

motion parallax



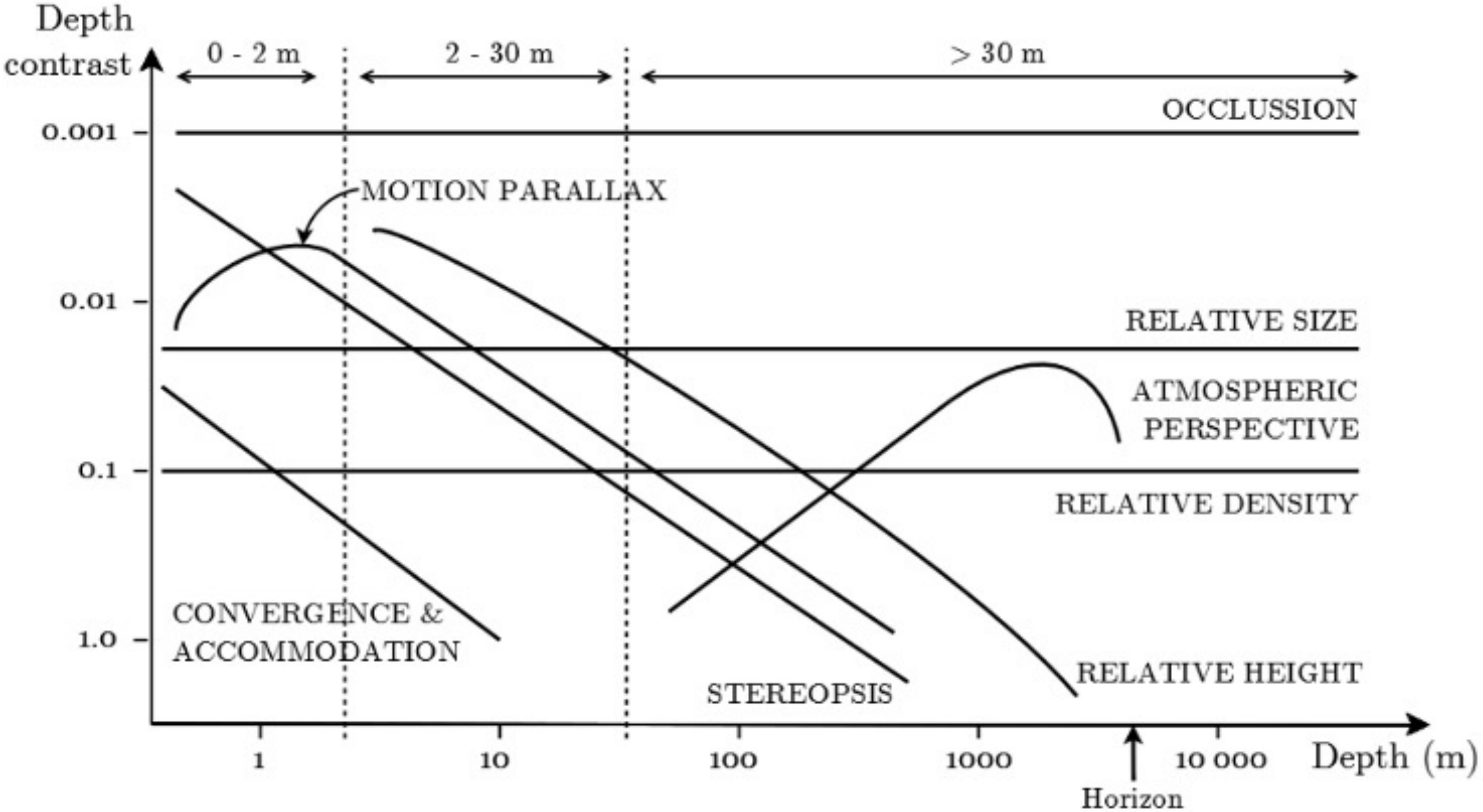
accommodation/blur

current glasses-based (stereoscopic) displays

near-term: light field displays

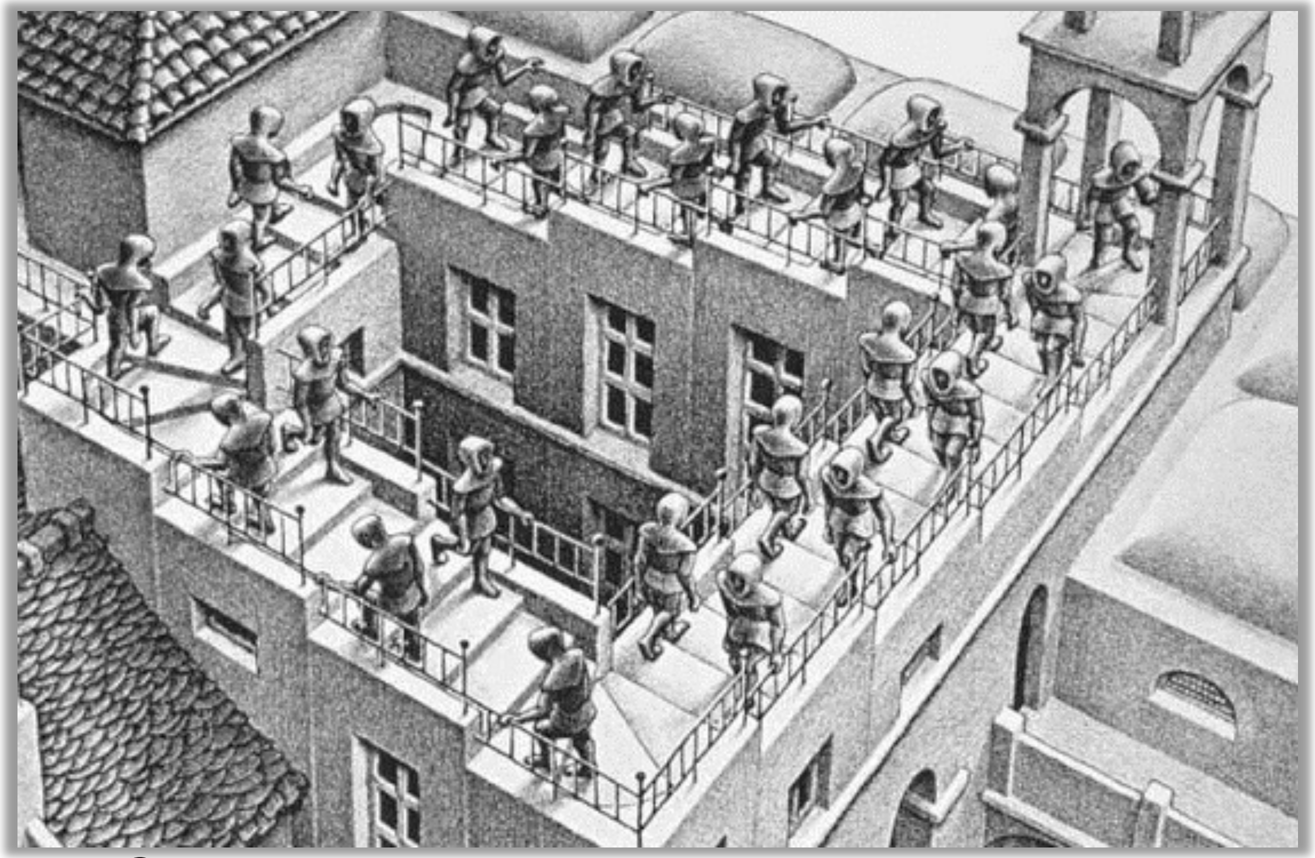
longer-term: holographic displays

Depth Perception

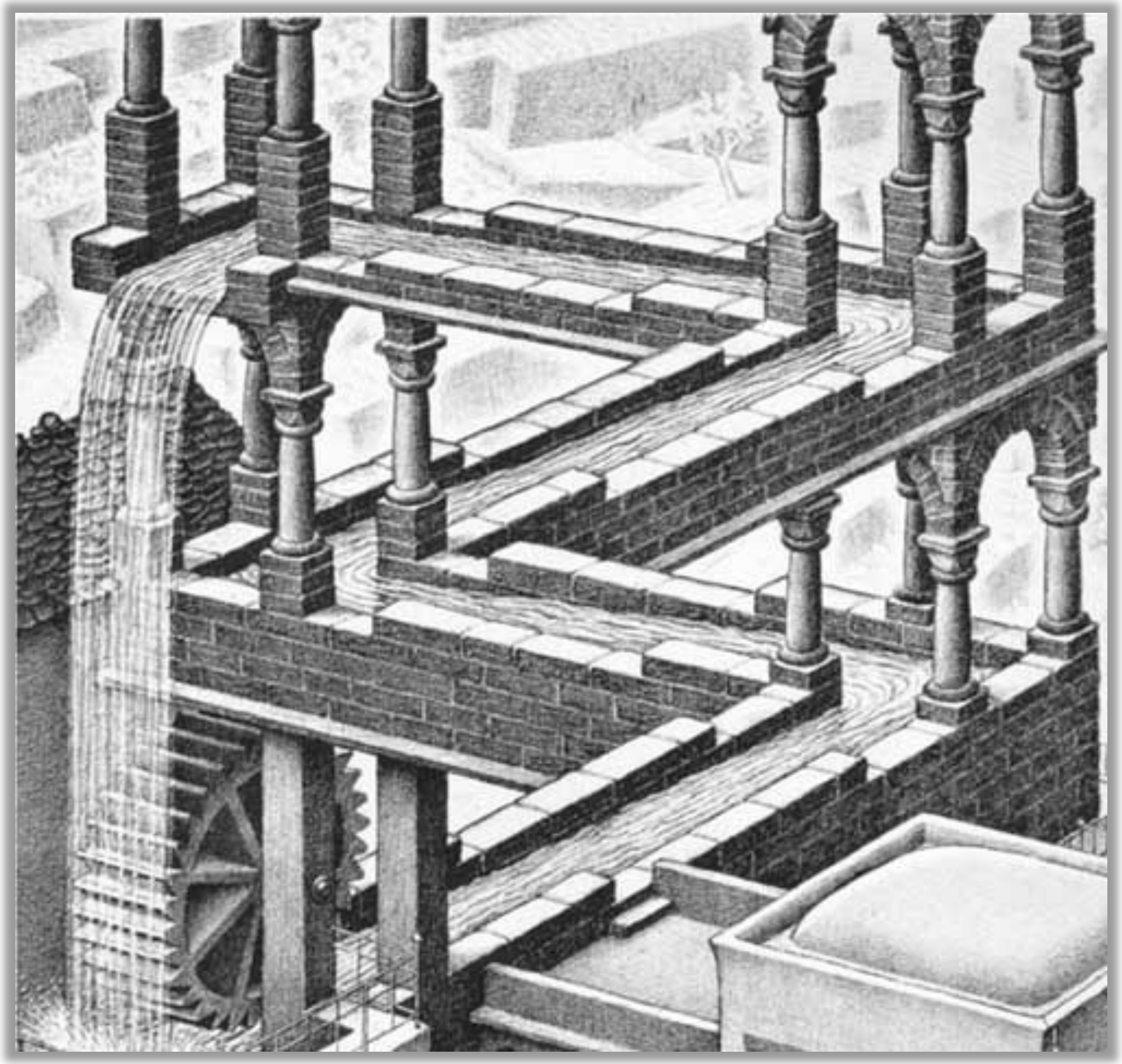


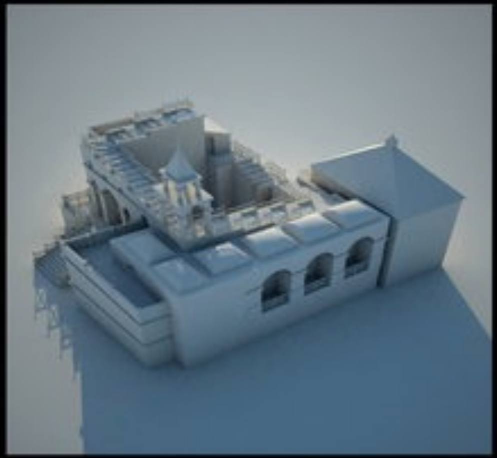
Cutting & Vishton, 1995

Visual Illusions – Perspective, Occlusion, Size



M.C. Escher

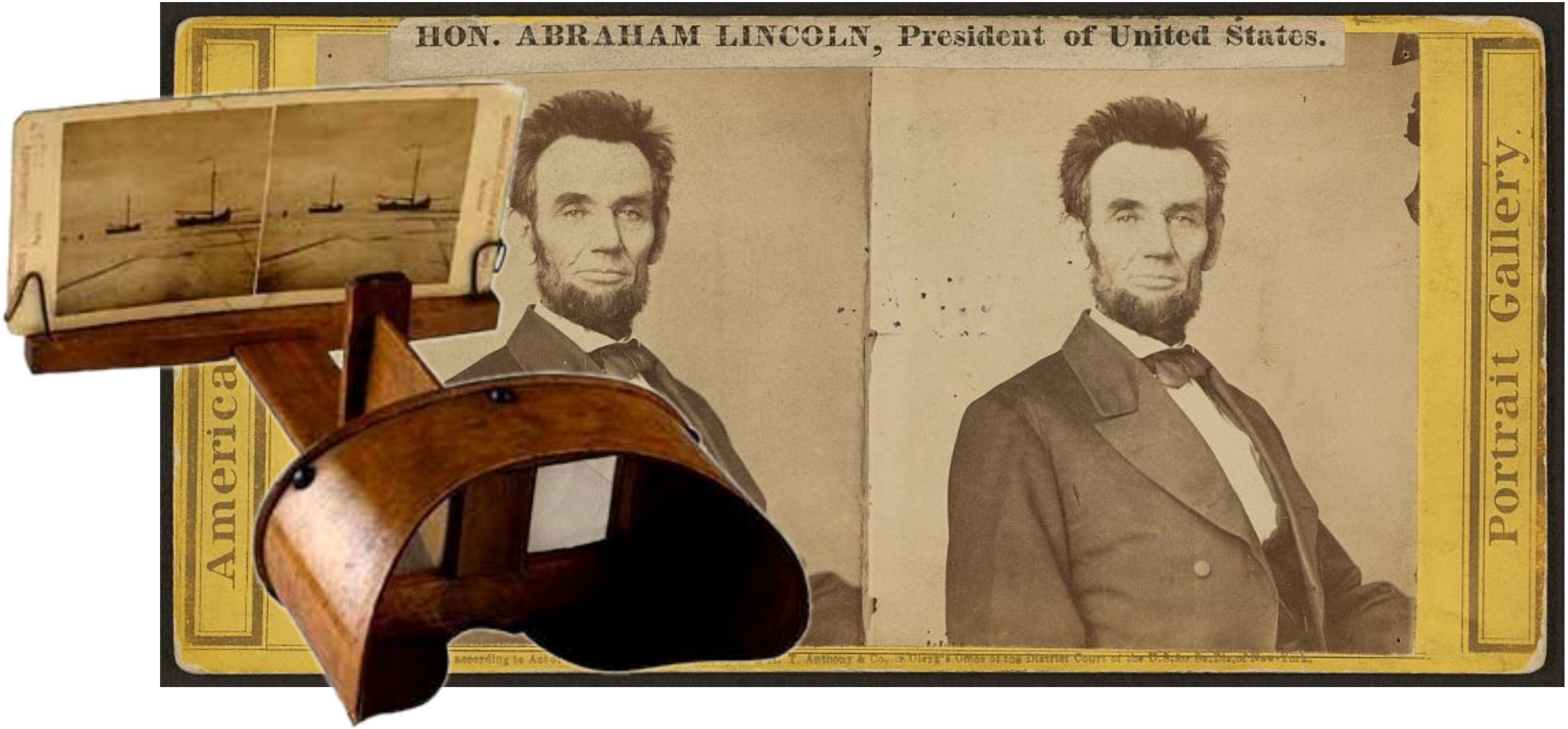


Visual Illusions – Which Cues are These?

Held et al., 2006, ACM SIGGRAPH

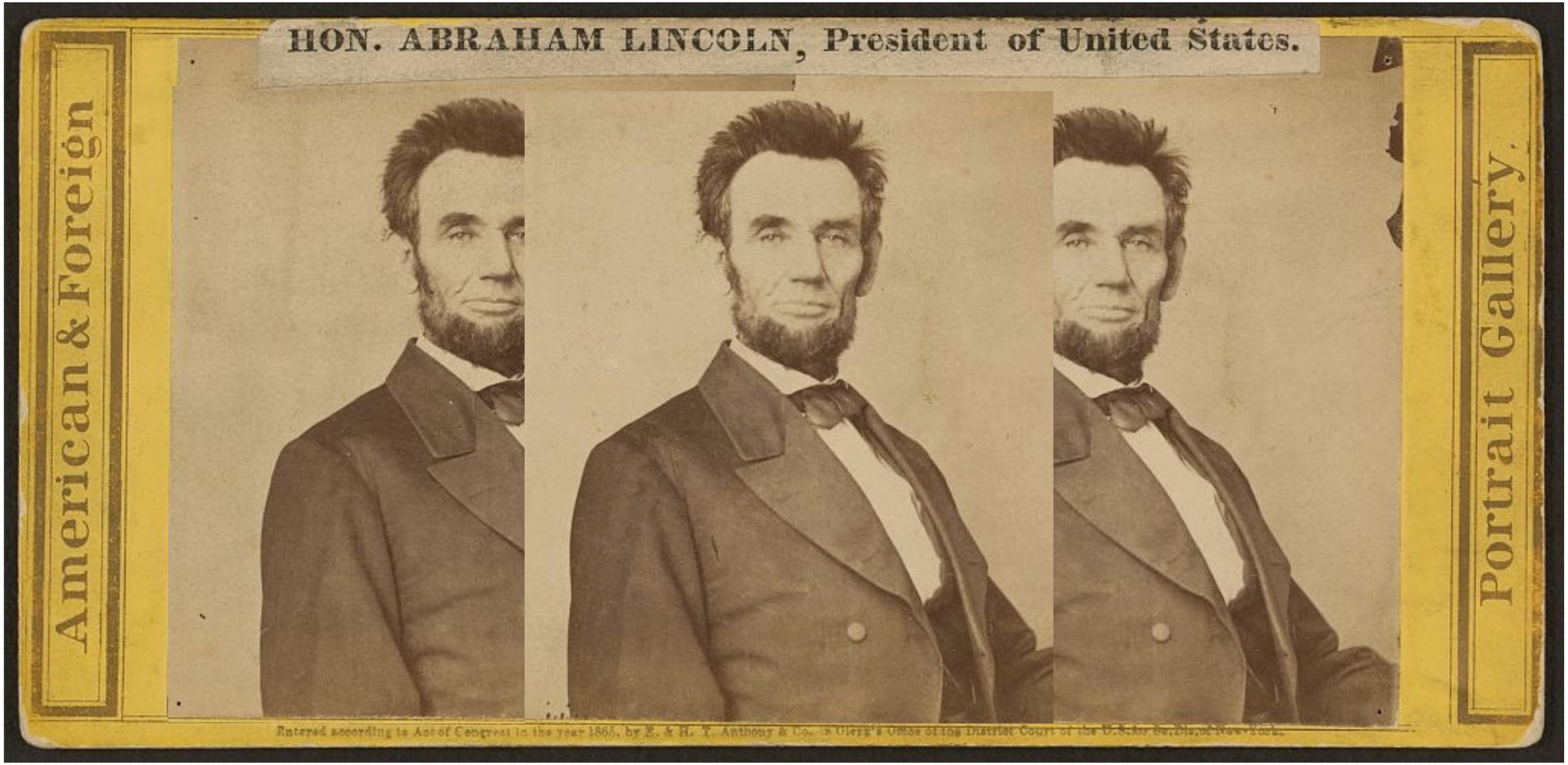
Stereoscopic Displays



Charles Wheatstone., 1841. Stereoscope.

Walker, Lewis E., 1865. Hon. Abraham Lincoln, President of the United States. Library of Congress

Stereoscopic Displays



Stereoscopic Displays

176 years later

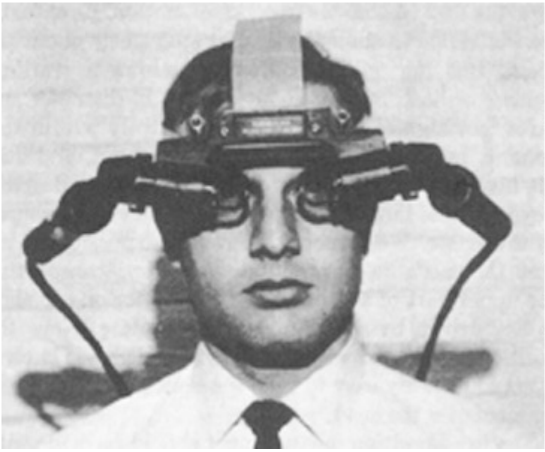
Charles Wheatstone 1838

stereoscopic displays

A Brief History of Virtual Reality

Stereoscopes
Wheatstone, Brewster, …

VR, AR,
Ivan Sutherland



VR explosion
Oculus, Sony, Valve, MS, …

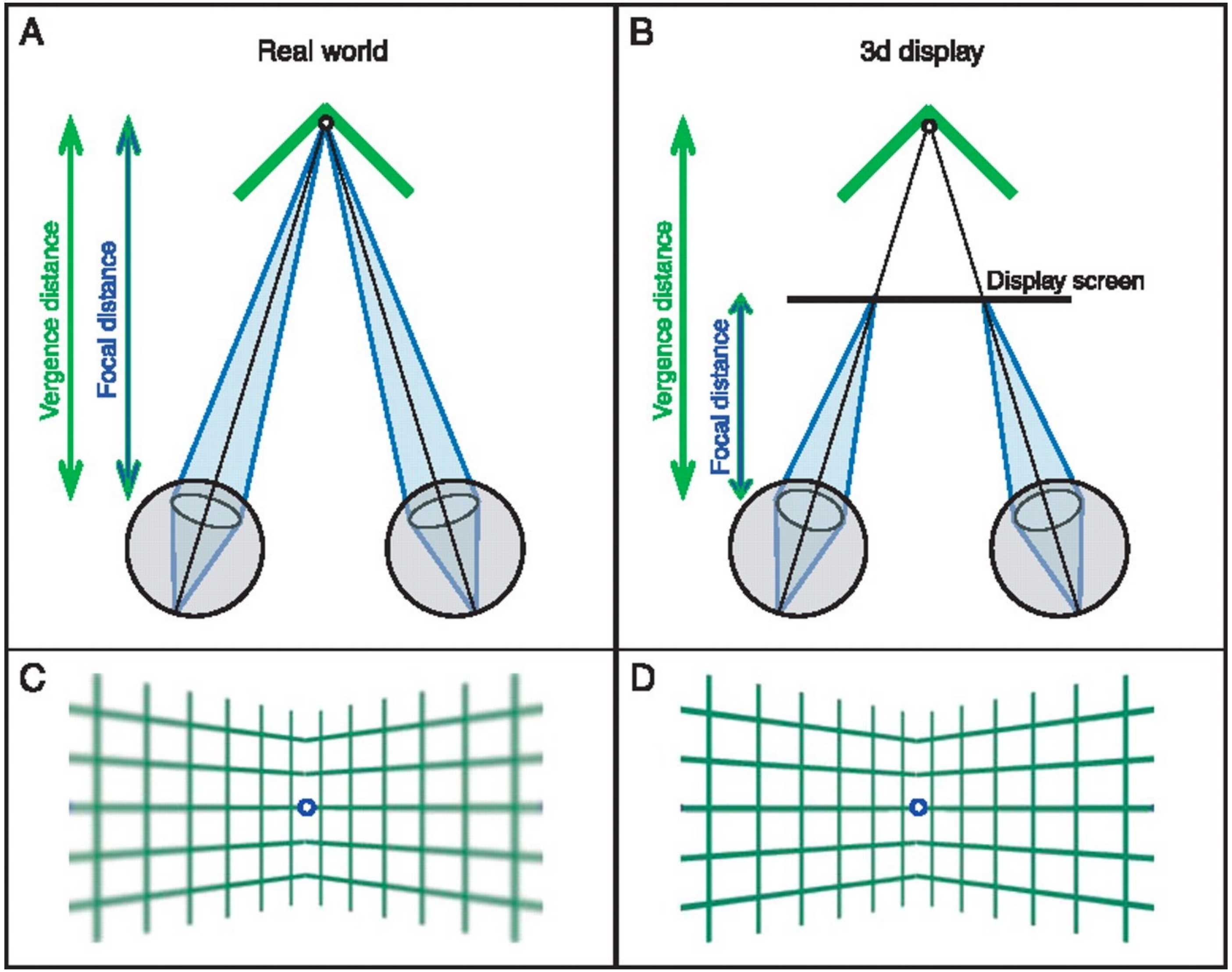
1838

1968

2012-2022

Next-generation VR/AR Displays

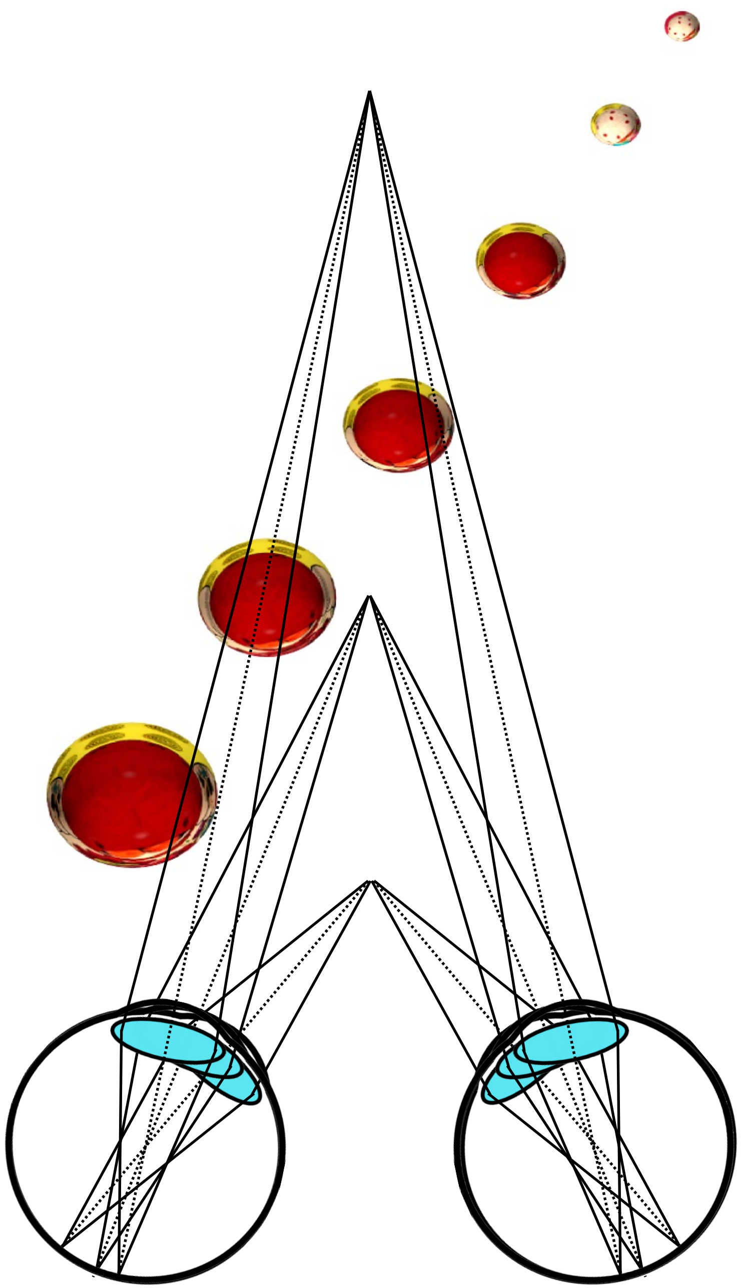
Vergence-Accommodation Conflict



Marty Banks, UC Berkeley

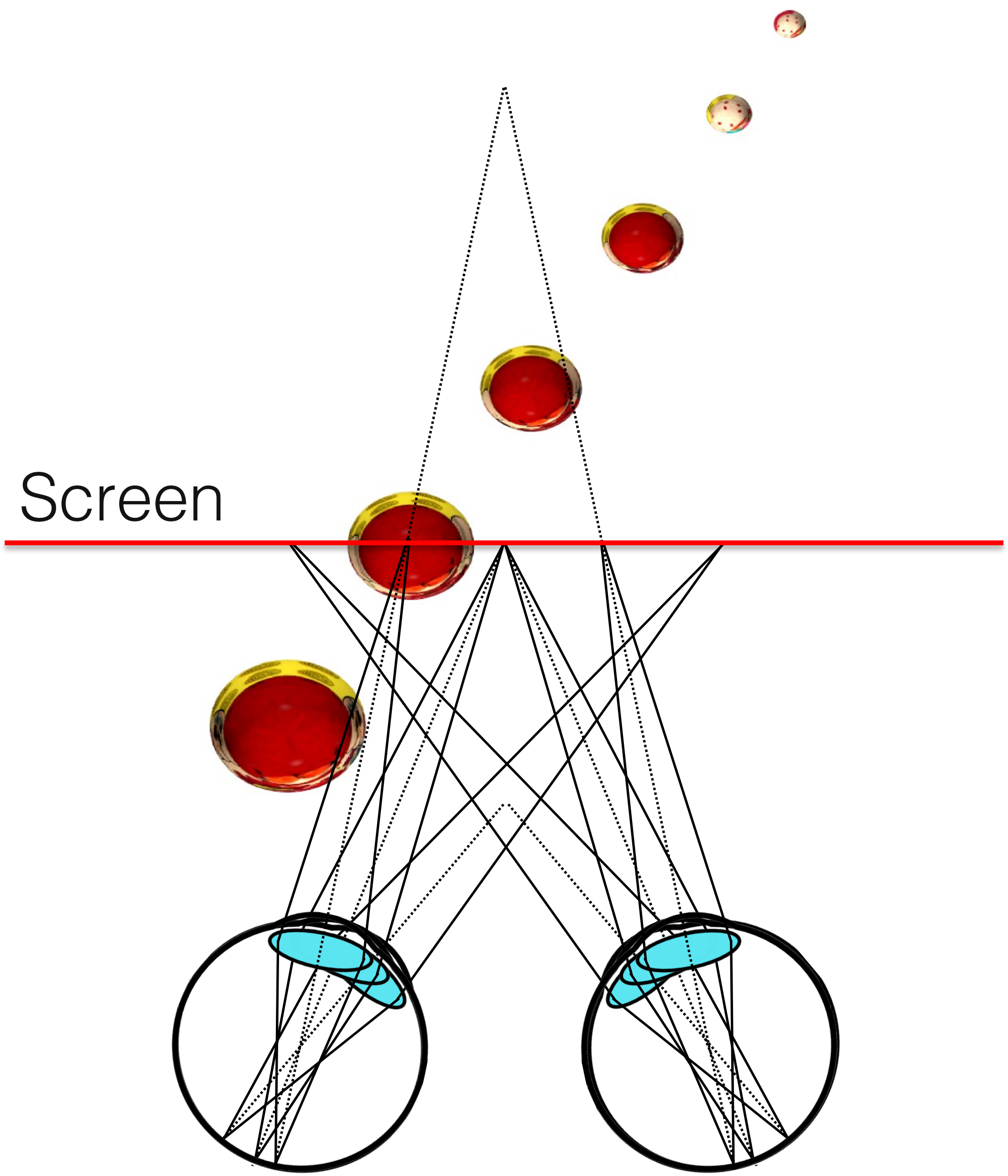
effects

- visual discomfort
- visual fatigue
- nausea
- diplopic vision
- eyestrain
- compromised image quality
- pathologies in developing visual system
- …



Real World:

Vergence & Accommodation Match!

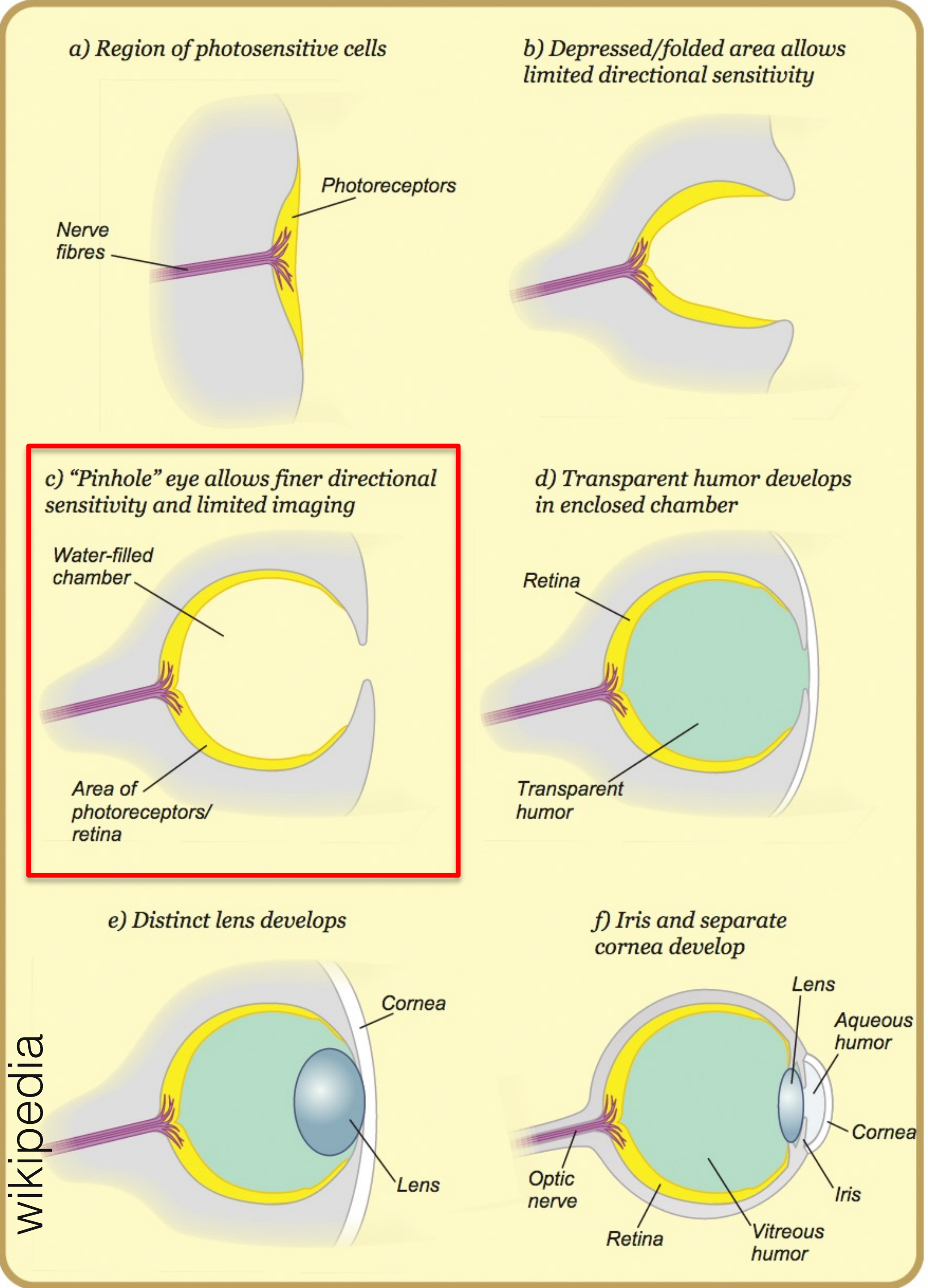


Stereo Displays Today:

Vergence-Accommodation Mismatch!

Summary

- visual acuity: 20/20 is ~1 arc min
- field of view: ~190° monocular, ~120° binocular, ~135° vertical
- temporal resolution: ~60 Hz (depends on contrast, luminance)
- dynamic range: instantaneous 6.5 f-stops, adapt to 46.5 f-stops
- color: everything in the CIE xy diagram; distances are linear in CIE Lab
- depth cues in 3D displays: vergence, focus, conflicts, (dis)comfort
- accommodation range: ~8cm to ∞, degrades with age



Homework I

- take a step back in evolution
- build a pinhole camera
- capture photos with it
- read instructions carefully!

Task 1 & 2: Create a pinhole camera

Pinhole

Camera lens
Goes here

Homework I – Build a Pinhole Camera

Next: Digital Photography I

- optics
- aperture
- depth of field
- field of view
- noise
- sensors
- color filter arrays

References and Further Reading

interesting textbooks on perception:

- Wandell, “Foundations of Vision”, Sinauer Associates,1995
- Howard, “Perceiving in Depth”, Oxford University Press, 2012

depth cues and more:

- Cutting & Vishton,” Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth”, Epstein and Rogers (Eds.), Perception of space and motion, 1995
- Held, Cooper, O’Brien, Banks, “Using Blur to Affect Perceived Distance and Size”, ACM Transactions on Graphics, 2010
- Hoffman and Banks, “Focus information is used to interpret binocular images”. Journal of Vision 10, 2010
- Hoffman, Girshick, Akeley, and Banks, “Vergence-accommodation conflicts hinder visual performance and cause visual fatigue”. Journal of Vision 8, 2008
- Huang, Chen, Wetzstein, “The Light Field Stereoscope”, ACM SIGGRAPH 2015

the retina and visual acuity:

- Roorda, Williams, “The arrangement of the three cone classes in the living human eye”, Nature, Vol 397, 1999
- Snellen chart: https://en.wikipedia.org/wiki/Snellen_chart

the visual field:

- Ruch and Fulton, Medical physiology and biophysics, 1960

contrast sensitivity function & hybrid images:

- Oliva, Torralba, Schyns, “Hybrid Images”, ACM Transactions on Graphics (SIGGRAPH), 2006
- Spatio-temporal CSF: Kelly, Motion and Vision. II. Stabilized spatio-temporal threshold surface, Journal of the Optical Society of America,1979
- Mantiuk, Kim, Rempel, Heidrich, “HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions”, SIGGRAPH 2011